4-26-2019

Improved Pion Correlation Functions Under Lattice Action

Daniel Hoying
University of Connecticut - Storrs, daniel.hoying@uconn.edu

Improved Pion Correlation Functions Under Lattice Action

Daniel Hoying, PhD, University of Connecticut, 2019

Abstract

We present the first calculation of pion-pion ($\pi\pi$) scattering at physical quark mass from lattice QCD. We simulate QCD in a periodic ~ 5 fm box at inverse lattice spacing $a^{-1} = 1.015, 1.3784$ GeV using zMöbius/Möbius Domain Wall Fermions and Iwasaki Gauge Action. We form operators composed of localized hydrogen-like wave functions for scalar ($\overline{\psi}\psi$), pseudoscalar ($\overline{\psi}\gamma_5\psi$), and vector bilinears ($\overline{\psi}\gamma_\mu\psi$) We then calculate all-to-all (A2A) meson field propagators for up to three units of individual particle momenta as well as isospin $I = 0, 1, 2$, with up to (in the $I = 0, 2$) three units of center of mass momentum. This allows us to project the resulting $O(10000)$ correlation functions onto definite isospin and irreducible representation (which allows us to project onto the lowest bose-symmetry allowed angular momenta of each of our operators). These projections form correlation function matrices of definite time separation, which then naturally define a generalized eigenvalue problem (GEVP) we can solve for the desired spectra. We then apply the Lüscher method to obtain from our lattice energies continuum scattering phase shifts. After partially accounting for some of our systematic errors, we obtain fair to good agreement with phenomenological predictions for the phase shifts as derived from Roy equations and chiral perturbation theory. These results can help serve as a foundation on which to build a fully periodic calculation of a kaon decaying into two pions ($K \to \pi\pi$), a very important decay related to the study of baryogenesis in the Standard Model. This work also contains algorithmic and mathematical developments which, while not used directly in the $\pi\pi$ scattering study, are likely to be useful in the study of $K \to \pi\pi$ (and might also be useful to interaction physics generally). These developments are highlighted briefly in the introduction.

Improved Pion Correlation Functions Under Lattice Action

Daniel Hoying

M.S., Ohio State University, 2014

B.S., Ohio State University, 2011

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

University of Connecticut

2019

APPROVAL PAGE

Doctor of Philosophy Dissertation

Improved Pion Correlation Functions Under Lattice Action

Presented by

Daniel Hoying, B.S., M.S.

Major Advisor _____
Thomas Blum

Major Advisor _____
Taku Izubuchi

Associate Advisor _____
Gerald Dunne

Associate Advisor _____
Luchang Jin

University of Connecticut
2019

Acknowledgements

I would like to thank my advisors, Prof. Tom Blum and Taku Izubuchi for their dedication and vision for the periodic project, their teaching and guidance, and their patience. I would also like to thank Chris Kelly for his large and continued contributions/advice to the project (and for teaching me a lot about programming). I'd also like to thank Chulwoo Jung for his help and encouragement. I'd like to thank Luchang Jin for his continued interest in the project and insightful suggestions. I'm grateful to Christoph Lehner, Norman Christ, Peter Boyle, Amarjit Soni, Aaron Meyer, David Murphy, Mattia Bruno, Jiqun Tu, and many others for their inspiring work solving the many problems we faced in this project.

Finally, I'd like to thank my friends and family for their continued support, especially my parents.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

We are motivated to calculate $\pi\pi$ correlators for several reasons. First, $\pi\pi$ correlation functions can be used as a direct comparison of many low energy QCD experiments vs. first principles theory. Second, we can study other physical processes which have some relationship with $\pi\pi$ scattering, $K \to \pi\pi$ being a prominent example and primary motivating force for this calculation. Third, we can use it to check calculations done using other lattice setups such as different boundary conditions (such as G-parity boundary conditions). This is related to $K \to \pi\pi$ since the existing calculation of $K \to \pi\pi$ on the lattice uses G-parity boundary conditions. Our collaborators also calculate (physical quark mass) $\pi\pi$ scattering (using different setups), so our periodic version can potentially inform their efforts. Finally, $\pi\pi$ scattering has not been studied using physical quark masses. Previous studies [1][2][3][4][5] used heavier quark masses to make the Dirac operator easier to invert. This introduces a non-perturbative systematic which is difficult to fully control without using physical quark mass.

Now we turn to $K \to \pi\pi$. The amount of CP violation (CPV) in $K \to \pi\pi$ decays is a possible explanation for matter/antimatter (M/AM) asymmetry in the Universe. Baryogenesis (which is another name for M/AM asymmetry), requires violation of CP. Furthermore, the amount of CP violation in the Standard Model appears too low to describe measured M/AM asymmetry giving us a hint of new physics.

Direct CPV was first observed in late 90s at CERN (NA31/NA48) and Fermilab (KTeV) in $K^0 \to \pi\pi$ [6]:

$$\eta_{00} = \frac{A(K_L \to \pi^0\pi^0)}{A(K_S \to \pi^0\pi^0)},\ \eta_{\pm} = \frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)}$$

$$Re(\frac{\epsilon'}{\epsilon}) = \frac{1}{6}\left(1 - \left|\frac{\eta_{00}}{\eta_{\pm}}\right|^2\right) = 1.66(23)\ x\ 10^{-3}\text{(Experiment)}$$

The ratio $\frac{\epsilon'}{\epsilon}$ defines the amount of direct CPV in K decays. We refer to this ratio colloquially as ϵ' (as this is the more difficult parameter to access).

In terms of isospin states, $\Delta I = 3/2$ decays to $I = 2$ final states, amplitude A_2

$\Delta I = 1/2$ decays to $I = 0$ final states, amplitude A_0

$$
\begin{aligned}
A(K^0 \to \pi^+\pi^-) &= \sqrt{\frac{2}{3}} A_0 e^{i\delta_0} + \sqrt{\frac{1}{3}} A_2 e^{i\delta_2} \\
A(K^0 \to \pi^0\pi^0) &= \sqrt{\frac{2}{3}} A_0 e^{i\delta_0} - 2\sqrt{\frac{1}{3}} A_2 e^{i\delta_2} \\
\Rightarrow \epsilon' &= \frac{i\omega e^{i(\delta_2-\delta_0)}}{\sqrt{2}} \left(\frac{Im\ A_2}{Re\ A_2} - \frac{Im\ A_0}{Re\ A_0} \right) \boxed{\omega = \frac{ReA_2}{ReA_0}}
\end{aligned} \tag{1.1}
$$

There is a large cancellation in the highlighed term in parentheses in eq. (1.1). This large cancellation taken with the small size of ϵ' makes it particularly sensitive to new direct-CPV introduced by most BSM models. The problem of calculating ϵ' is further reviewed in many sources. For a thorough introduction, see (for instance) [7].

Because our main goal is to calculate scattering phase shifts, we can divide the thesis into parts which directly facilitate that goal and parts which do not. To outline the former, we first illustrate the steps involved in computing phase shifts: We use existing periodic ensembles to compute eigenvectors (section 2.0.1,[8]) and gauge fixing matrices (see, e.g., [9]). On many ensembles, we calculate high modes via deflated conjugate gradient (split-CG section 2.3); deflation from the eigenvectors). We then gauge fix to Coulomb gauge (see explanation in section 3.2). Once we've inverted the Dirac operator, we compute meson fields (again, see section 2.0.1), then correlation functions section 1.1. At this point, we've finished production on the supercomputer. The remaining post-processing and analysis can be done on a personal computer. We begin post-processing by vacuum subtracting our $I = 0$ correlators (see chapter 6). We then jackknife to compute errors eq. (2.10), project onto definite isospin (chapters 4 and 5)/irrep (chapter 7;[10][1]). In order to remove an important systematic, we must then subtract around the world terms (see section 8.4). Once complete, we move to extracting $\pi\pi$ lattice spectra by solving a generalized eigenvalue problem (GEVP) (see chapter 8). After solving the GEVP, we find effective masses (section 9.2). At this point we in principle have our interacting $E_{\pi\pi}$ (albeit they are dependent on time). However, we apply the pion ratio method (section 8.10) to further reduce statistical errors. We extract final energies via fits to the time slice effective mass data. From here, we can compute phase shifts via Lüscher ([11]) formalism (see further section 1.3.1,[12]). Lastly, we check to see if we've accomplished our goal (see section 10.7) by comparing to chiral perturbation theory [13] as well as phenomenology extrapolated from experimental data [14].

Besides being the first study of pions with physical quark mass, this thesis presents other contributions: First, we find a new, general symmetry of lattice correlators that can be used to reduce correlator combinatorics (section 2.1). Second, we present an improved way to subtract around the world systematic error (sections 8.4.2, 8.5 and 10.8). Third, we present a systematic study of different ways to perform vacuum subtraction (in the $I = 0$ $\pi\pi$ sector) (chapter 6). Fourth, we examine the excited state perturbation theory of 2pt correlation functions and offer alternative ways to extract excited states (as well as a possible[1] improvement over the existing convergence rate) (chapter 8). Fifth, we also examine the examination of the excited state perturbation theory of 3pt correlation functions, which offer a possible improved convergence rate for the extraction of matrix elements (crefGammaref). Sixth, we show a new, general statistical technique to handle fit range systematics (p-value weighted fit range averaging: section 9.3). We are also among the first to use new state of the art techniques to reduce noise in pion correlators including sections 8.7.1 and 8.10. Seventh, we find a breadth-first search algorithm based on the statistical properties of the χ^2 distribution which allows for fast looping over possible fit ranges ([15],item 2). Eighth, we have a method to perform A2A subtraction when using the MADWF procedure (section 1.3.2.1). Finally, we present an improved communication strategy for contractions involving many (node-distributed) meson fields (section 2.2.1).

We now outline the thesis chapters. In the remainder of chapter 1 we present lattice QCD background. In chapter 2 we have information about computation techniques and strategies especially as they relate to production performance. In chapter 3, we present details about the gauge ensembles we used to computed correlation functions. In chapters 4 and 5 we present wick contractions which constitute the isospin projection we apply to our correlation functions. In chapter 6, we explain our method to remove zero energy contributions to our $I = 0$ correlation functions. In chapter 7, we explain angular momentum on the lattice, and our angular momentum projections. chapter 8 has extensive explorations and explanations of our GEVP techniques and possible extensions of this framework with the goal of eliminating excited state contamination. chapter 9 describes our fitting techniques (which are also available in code form from [15]). In chapter 10, we present results from our calculation including phase shift plots (numerical results are in chapter A). chapters 11 and 12 has, respectively, $K \to \pi\pi$ and $K \to \sigma$ contractions which are needed for the planned $K \to \pi\pi$ periodic project. The final chapter, chapter 13 has conclusions and future work planned in this line of research.

The appendices have technical production details, numerical results, consistency checks, and other miscellaneous notes. They also have all the effective mass plots (see chapter A) which generate the summarizing graphs of chapter 10.

[1] These methods still need to be numerically tested.

1.1 Lattice QCD - Overview

We start with a simple path integral in Minkowski space.

$$\left\langle x_f | e^{-iHt} | x_i \right\rangle = \int \mathscr{D}x(t) e^{-iS[x]} \tag{1.2}$$

We know that the computer can only represent real numbers in finite precision, and we might get arbitrarily large cancellations of phases if we try to evaluate eq. (1.2) computationally. This is referred to as the sign problem. Hence, we take $t \to -it$ (and work at zero temperature and zero chemical potential). We also know that when we move from quantum mechanics to full field theory we replace $x \to \phi(x)$ (the field is the analogue of the position). We thus end up with the path integral integrated over fields

$$\begin{aligned} \langle \Gamma[\phi] \rangle &\equiv \frac{1}{Z} \int e^{-S[\phi]} \Gamma[\phi] \prod_{x_j} d\phi(x_j) \\ Z &\equiv \int e^{-S[\phi]} \prod_{x_j} d\phi(x_j) \end{aligned}$$

where Γ is some observable. Observables we might calculate on the lattice are diverse. We usually calculate N-point correlation functions (where N is the number of time slices with an explicit operator insertion. The most common object we calculate on the lattice is a 2-point function:

$$\begin{aligned} \left\langle 0 | O_i(t_{\text{tsnk}}) O_j^\dagger(t_{\text{tsrc}}) | 0 \right\rangle &= \left\langle 0 | e^{-Ht_{snk}} O_i e^{Ht_{\text{snk}}} e^{-Ht_{src}} O_j^\dagger e^{Ht_{\text{src}}} | 0 \right\rangle \\ &= \left\langle 0 | O_i e^{-H(t_{src} - t_{\text{snk}})} O_j^\dagger | 0 \right\rangle \end{aligned}$$

where e^{-Ht} is our time evolution operator, $|0\rangle$ is the vacuum, O_i is our destruction operator at the sink and $O_j^\dagger$ is our creation operator at the source. If we insert a complete set of states of (eigenvectors of H) $\sum_n^\infty |n\rangle \langle n|$, we can see that our correlation functions decay exponentially with time:

$$\begin{aligned} &= \sum_n^\infty \left\langle 0 | O_i e^{-H(t_{src} - t_{\text{snk}})} | n \right\rangle \langle n | O_j^\dagger | 0 \rangle \\ &= \sum_n^\infty e^{-E_n(t_{src} - t_{\text{snk}})} \left\langle 0 | O_i | n \right\rangle \langle n | O_j^\dagger | 0 \rangle \end{aligned}$$

Notice also that, the poles and cuts we used to have in our scattering problem are replaced by decaying exponentials. In general, we do not know the eigenstates of our Hamiltonian, so states we construct on the

lattice have excited state contamination. This contamination decays exponentially, but signal to noise also decays exponentially with time (decaying faster for higher energies). This implies that we must find some ways to eliminate contamination from states which we don't have the data to resolve (see chapter 8).

Our $\pi\pi \to X$ processes (where $X \in \{\pi\pi, \sigma, \rho\}$) involve two point functions. $K \to \pi\pi$ needs a three point correlation function as we insert an operator O_w in between our source and sink operator.

1.2 Continuum Limit

Lattice QCD is an effective field theory, meaning we would like to integrate out degrees of freedom (operators) associated with length scales shorter than our lattice spacing a (and energies higher than a^{-1}) by using perturbation theory. These higher dimension operators are part of the full action, so if we don't include them somehow (e.g. via perturbation theory) we get discretization errors.

Our lattice Lagrangian is

$$\mathscr{L}_L = \mathscr{L}_c + a\mathscr{L}_5 + a^2\mathscr{L}_6 + O(a^3)$$

where $\mathscr{L}_c$ is the continuum part of the lagrangian, $\mathscr{L}_5$ is a dimension 5 operator proportional to $m_{\text{ res}}$ which breaks chiral symmetry (such as $\overline{q}\sigma_{\mu\nu}F^{\mu\nu}q$). In the domain wall action, we eliminate (or exponentially suppress in the size of the fifth dimension) this chiral symmetry breaking part of the Lagrangian. The so-called Symanzik-improved action eliminates this term by tuning these higher dimension operators for Wilson fermions. We thus see that (in the DWF case) we have $O(a^2)$ discretization error.

We can address $O(a^2)$ error via perturbation theory, but a better solution is to compute at several lattice spacings and extrapolate to the continuum limit assuming we can neglect terms smaller than $O(a^3)$, as $O((a\Lambda_{QCD})^2)$ is anyway $\sim 10^{-2}$. If we have two measurements of an observable $\langle O\rangle_{\text{ coarse}}, \langle O\rangle_{\text{ fine}}$, then we can find $\langle O\rangle_{\text{ continuum}} \equiv \langle O\rangle$ via the solution to

$$\langle O\rangle = \langle O\rangle_{\text{ coarse}} + \omega a^2_{\text{coarse}}$$
$$\langle O\rangle = \langle O\rangle_{\text{ fine}} + \omega a^2_{\text{fine}}$$

where ω is our scaling parameter.

1.3 Finite Volume Effects

We have errors associated with our infrared cutoff, the box size L. These errors for single particles should fall off as $e^{-m_\pi L}$ due to well-known arguments of Lüscher [16] and for two particles scattering should receive power law corrections in $\frac{1}{L}$ [17], but it turns out we can avoid these power law corrections if we use the method from section 1.3.1.

We also know the finite volume gives us a quantization condition we can impose on our spectrum. This quantization condition is independent of the lattice spacing, meaning it can relate continuum and lattice quantities. In fact, if we know the continuum infinite volume energy spectrum, we can impose our quantization condition and find the allowed finite volume energies. If we also know the scattering phase shifts from the continuum energies, we can relate these phase shifts to the allowed energy levels (allowed energy eigenstates of our finite volume Hamiltonian).

In other words, assuming we measure the lattice spectrum, we can reverse this whole process and get the continuum scattering phase shifts up to discretization error mentioned in section 1.2. Thus, it is a bit of a misnomer (in the usual exposition on this topic) to say we can get the continuum phase shifts from the lattice energies. Rather, we can get continuum phase shifts if we have continuum energies, and we can get which of these should be measurable on our lattice from our quantization condition.

1.3.1 Setting up the Scattering Problem

[12]

We work in the center of mass frame (CMF). We know we have some potential dependent on the relative distance between our two particles $V(x)$ which we suppose goes to 0 well before we reach the boundary of our box (or, in other words, the interaction region is smaller than the lattice size). That is, our interaction should be short range (like QCD). Outside the box, the particles propagate without interaction so we have

$$\left(\nabla^2 + p^{*2}\right)\phi(x) = 0$$

which is the well-known Helmholtz equation. The general solution to this equation in terms of in-going and out-going partial l-waves from our interaction region has parameters which fully define the phase shift. We match this at the boundary to our interior region wave function $\psi(x_1, x_2)$ where x_1, x_2 are the positions of the

two particles which obeys

$$\psi(x_1, x_2) = \psi(x_1 + n_1 L, x_2 + n_2 L) \tag{1.3}$$
$$n_1, n_2 \in \mathbb{Z}$$

eq. (1.3) is our quantization condition. The details are omitted, but we finally derive, via this general expansion, a relationship between δ_l the scattering phase shift of angular momentum l and our finite volume energies (which fully determine the parameter q^2):

$$det\left[e^{2i\delta}(M-i)-(M+i)\right] = 0 \tag{1.4}$$
$$M_{lm,l'm'} \equiv M_{lm,l'm'}(q^2)$$
$$\left[e^{2i\delta}\right]_{lm,l'm'} \equiv e^{2i\delta_l(p^*)}\delta_{ll'}\delta_{mm'}$$

Further details including a derivation can be found in, e.g., [12].

1.3.1.1 Boosted Lüscher Zeta $\mathcal{Z}_{00}$

The quantization condition eq. (1.4) can be written (for the lowest spins $J = 0, 1$) as in the (in general) boosted frame as [11][18][19]:

$$\tan\delta(\overline{p}_n) = \frac{\gamma\pi^{3/2}\sqrt{\overline{m}}}{\mathcal{Z}_{00}(1;\overline{m})}$$
$$\mathcal{Z}_{00}(s,\overline{m}) = \frac{1}{\sqrt{4\pi}}\sum_{\mathbf{r}\in P}(r^2-\overline{m})^{-s}$$
$$P \equiv \{\mathbf{r}|\mathbf{r} = \hat{\gamma}^{-1}\left(\mathbf{n}+\frac{\mathbf{d}}{2}\right), \mathbf{n}\in\mathbb{Z}^3\}$$

Details for calculating $\mathcal{Z}_{00}$ ($\mathcal{Z}_{lm}$ is the generalized Lüscher zeta function; for all of our phase shifts we have $l = 0, m = 0$) can be found in, e.g., [20]. Thus, our phase shift is a function only of $\overline{p}_n$ where we measure on the lattice center of mass energy E_{cm} and

$$E_{cm} = 2\sqrt{m_\pi^2 + \overline{p}_n^2}$$

1.3.2 Möbius/zMöbius Domain Wall Fermion Kernels

In this section, we very briefly touch on the ideas of Kaplan [21] and Shamir [22] concerning Domain Wall Fermions (DWF) and their generalizations. rational polynomial First, in a naive QCD simulation in $3+1$

dimensional box (in fact, any theory of odd spatial dimensions), we generically have the problem of doublers. Suppose we create a particle with n units of lattice momentum in one direction. Naively, our momentum would be $p_n = \frac{2\pi n}{L}$. In momentum space we must have periodicity so $n + L = n$. Then our momentum must actually be $hatp_n = sin(\frac{2\pi n}{L})$. Then our dispersion relation becomes

$$\hat{p}_n^2 = E_n^2 - m^2$$

That is, for every solution n, we get another solution. This doubling happens for every dimension of our problem, and the partners come in opposite chiralities. In fact, this problem is quite generic for any theory of odd spatial dimension. This observation was proved by Nielsen and Ninomiya as the celebrated Nielsen-Ninomiya no-go theorem [23]. An elegant (and much less difficult to parse) proof can be found in [24].

We can add a Wilson term which is a bilinear with a second (covariant) derivative which gives all of the doublers except one a very large mass which causes these particles to decouple. Unfortunately, this badly breaks chiral symmetry. We can add dimension 5 operators and tune the coefficients to elimiante the (on-shell) $O(a)$ chiral symmetry breaking, but this introduces other complications (such as off-shell chiral symmetry breaking) and the tuning in general needs to be quite precise. One way to get around the no-go theorem (first described in [21]) which preserves chiral symmetry (up to small residual mass term m_{res}) is to solve a theory in $4+1$ dimensions. We add a mass term which is a step function in the fifth dimension (which we denote by coordinate s). Then, only one of the (massless) chiral solutions is normalizable and it is bound to a $4d$ hyperplane at $s = 0$ with exponentially small extent into the fifth dimension. The other massless chiral mode is banished to the other end of the fifth dimension. The degree to which these small exponential tails overlap gives us our remaining chiral symmetry breaking, which we can eliminate as the size of the fifth dimension $L_s \to \infty$. The so-called domain wall mass term is set to be a step function $\epsilon_{L_s} \equiv [x]\frac{x}{|x|}$ in the fifth dimension (also referred to in the literature as the sign function [25]). The parameter x represents the eigenvalue spectrum of the H_5, the 5D DWF kernel [25]. It is approximated by:

$$\epsilon_{L_s}[x] = \frac{(x+1)^{L_s} - (x-1)^{L_s}}{(x+1)^{L_s} + (x-1)^{L_s}}$$

The sign function (technically $1-\epsilon^2$[25]) determines the amount of chiral symmetry violation. Additionally, we know that larger L_s, while decreasing m_{res} (and thus perfect chiral symmetry), increases computational costs linearly in its extent. We thus are motivated to form rational approximations of the sign function which are lower order polynomials in numerator and denominator, allowing us to reduce the size of L_s and save on

computation time. One can derive from these approximations coefficients for H_5 by finding the roots of $1-\epsilon^2$. These (quark mass independent) coefficients make ϵ^2 as close to 1 as possible for each s-slice along the 5th dimension (one has a nearly diagonal operator in the 5th dimension, so the coefficients are trivially related to the roots we find). If these coefficients (usually called the b,c coefficients, or ω parameters) are not constant with respect to s, we call the H_5 kernel the zMöbius kernel (independence of s gives Möbius) ([26]):

$$\epsilon_{L'_s,approx}[x] = \frac{\prod_s^{L'_s}(1+\omega_s^{-1}x) - \prod_s^{L'_s}(1-\omega_s^{-1}x)}{\prod_s^{L'_s}(1+\omega_s^{-1}x) + \prod_s^{L'_s}(1-\omega_s^{-1}x)}$$

$$b_s + c_s = 1/\ddot{\mathrm{I}}\mathrm{L}'_s$$

It turns out that because the zMöbius Dirac operator only has an approximate correspondence with the overlap operator, it breaks 5d γ_5 hermiticity (which is defined as $\gamma_5 M^\dagger \gamma_5 = M$ for some operator M). Thus, we must be correct the result via AMA procedure (see section 2.6).

This AMA procedure (see section 2.6 for an explanation of AMA) is known as Möbius accelerated domain wall fermions (MADWF). This solver for this kernel is defect correcting, and crucially is deflated using the low-modes of the zMöbius kernel. One transforms between Dirac operators of different L_s by projecting to and from the 4D overlap representation. For details, see [27].

1.3.2.1 4D A2A Subtraction

In the case of MADWF, our procedure is to first compose the low modes of the quark propagator from the zMöbius eigenvectors and eigenvalues. One usually performs A2A subtraction (see section 2.0.1) in the even-odd preconditioned matrix (eq. (2.12)) since the eigenvectors are 5D. However, the result of the MADWF procedure is to find solutions to the inverse of a different 5D operator with a larger L_s. Consequently, the usual subtraction procedure is not compatible. A very easy solution to this problem (which one can show is exact) is to wait until we have done the projection from MADWF to overlap (5D→4D) and after we've done a similar projection on the low modes from the zMöbius 5D vector space to the overlap vector space. We then can do the subtraction in this 4D space using A2A $|V\rangle, |W\rangle$ vectors.

Chapter 2

Computational Techniques

2.0.1 All-to-all Quark Quark Propagators (A2A)

We would like to invert the Dirac operator $\not{D}$ and eventually take the full volume average of the resulting correlation functions (via Fourier transform), but we know generically that finding $\sum_{x,y} e^{i\vec{y}\cdot\vec{p}_y} e^{-i\vec{x}\cdot\vec{p}_x} \not{D}^{-1}_{xy} \equiv \langle y|\not{D}^{-1}|x\rangle$ is an $O(V^2)$ operation in volume sums over x, y. In general, $O(V^2)$ operations on the lattice are prohibitively expensive. We would thus like to reduce the cost of this inversion. To do so, consider defining random noise vectors $|\eta\rangle$ with the property

$$\begin{aligned}
\delta_{xy} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} |\eta_i\rangle_x \langle\eta_i|_y \\
\not{D}^{-1}_{xy} &= \not{D}^{-1}_{xz} \delta_{zy} \\
\Rightarrow \not{D}^{-1}_{xy} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} \not{D}^{-1}_{xz} |\eta_i\rangle_z \langle\eta_i|_y
\end{aligned}$$

In practice, we use random $U(1)$ noise. We refer to N as the number of hits. Additionally, we would like to deflate this matrix inversion with the low modes of the Dirac operator ($\not{D}|v\rangle = \lambda_v |v\rangle$) which we calculate using the Lanczos algorithm. Deflation refers to the practice of projecting out the low modes of the Dirac operator. Since propagators generally are dominated by low modes, the amount of work we have to do (the matrix multiplication iteration count of our solver, the Conjugate Gradient (CG) algorithm) is greatly reduced after deflation. Thus, the full expression for our quark propagator becomes

$$\not{D}^{-1}_{xy,A2A}(N) \equiv \sum_v \frac{1}{\lambda_v} |v\rangle\langle v| + \frac{1}{N}\sum_{i=0}^{N}\left[\not{D}^{-1}_{xz}|\eta_i\rangle_z \langle\eta_i|_y - \sum_v \frac{1}{\lambda_v}|v\rangle_x \langle v|\eta_i\rangle\langle\eta_i|_y\right] \tag{2.1}$$

$$\not{D}^{-1}_{xy} = \lim_{N\to\infty} \not{D}^{-1}_{xy,A2A}(N)$$

The last term with the minus sign in eq. (2.1) we refer to as the A2A subtraction term.

We can dilute the $|\eta\rangle$ noise vectors so that the noise is confined to only a subset of the total mode space. In our case, we dilute in all but the spatial dimensions, so that our noise vectors are delta functions in all indices except space. The spatial part of the quark propagator is thus calculated stochastically, while the other degrees of freedom are calculated exactly.

$$|\eta_i\rangle_{s,c,t,x} = \delta_{ss'}\delta_{cc'}\delta_{tt'}|\eta_i\rangle_x$$

$$\delta_{xy} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N}|\eta_i\rangle_x\langle\eta_i|_y$$

We use hits to refer to only the spatial noise, so that for a single hit we end up inverting on

$$N_\eta = NN_cN_sL_t$$

noise vectors. For our lattices, where $L_t = 64, N_c = 3, N_s = 4$ we fix $N = 1$ and so have $N_\eta = 768$, which we refer to as the high modes.

We refer to the A2A sources as $|W\rangle$ and sinks $D^{-1}|W\rangle = |V\rangle$ so that

$$\not{D}^{-1}_{A2A} = \sum_i |V_i\rangle\langle W_i|$$

As it is usual impractical to save such high dimensional objects as $|V\rangle, |W\rangle$ and since we will need extensive computer resources to form inner products with some Γ structure and trace, we are not able to save the quark propagators as one might attempt in other schemes. We thus first take the inner product of

$$\langle V|\Gamma|W\rangle \tag{2.2}$$

With proper Fourier transforms, and with hydrogen-like wave function which connects the coordinates of $|W\rangle$ with $\langle V|$, we refer to these objects as meson fields (MF). We represent a pion on the lattice as a meson field (where $|V\rangle$ and $|W\rangle$ are the ends of a quark and anti-quark propogator). For a pion we have $\langle V|\Gamma|W\rangle$ and the Γ has some spatial structure we fix to

$$\Gamma(x,y) \propto \gamma^5 e^{i\vec{p}_x\vec{x}} e^{-i\vec{p}_y\vec{y}} e^{-|\vec{x}-\vec{y}|/r}$$

where r is the radius of our hydrogen-like wave function. We similarly project $|W\rangle, |V\rangle$ into momentum space via FFT. Projecting into momentum space allows us to more easily project onto energy eigenstates as well as implictly taking a full volume average of our correlation function.

The remaining uncontracted indices (which will be subsequently contracted) are the mode indices. Properly matched, we can then multiply meson fields together and take the trace over the remaining mode indices. We then save the results of these traces for every possible time combination (up to a fixed time separation relationship between the two particles of a given two particle operator).

2.0.1.1 Additional Hits - Theory

While ideally one could calculate all the hits at once, memory limitations often force us to have $N = 1$ for our initial run. If the gauge noise and the A2A noise converge at the same rate, we don't gain in statistical precision by rerunning on the same gauge configuration with a different set of $|\eta_i\rangle$. This has been seen to be true for the $\pi\pi$ sources, for instance. However, if our source has strong overlap with the high modes, our precision might benefit from an additional hit.

We should be concerned, however, about the normalization. We have low modes and high modes in our quark propagator, and we are counting the low modes twice whenever we rerun for an additional hit (because we use the same low modes but change the high mode source). We can see this by looking at a given correlation function:

$$Tr(...(\not{D}_l^{-1} + \not{D}_h^{-1})\Gamma(\not{D}_l^{-1} + \not{D}_h^{-1})...) = Tr(...\not{D}_l^{-1}\Gamma\not{D}_h^{-1} + \Gamma_{l\to h} + \Gamma_{h\to l} + \Gamma_{h\to h}...)$$

Now, we look at an additional hit.

$$Tr(...(\not{D}_l^{-1}+\frac{1}{2}\not{D}_{h_1}^{-1}+\frac{1}{2}\not{D}_{h_2}^{-1})\Gamma(\not{D}_l^{-1}+\frac{1}{2}\not{D}_{h_1}^{-1}+\frac{1}{2}\not{D}_{h_2}^{-1})...) =$$
$$Tr\left(...\Gamma_{l\to l}+\frac{1}{2}\Gamma_{l\to h_1}+\frac{1}{2}\Gamma_{h_1\to l}+\frac{1}{2}\Gamma_{l\to h_2}+\frac{1}{2}\Gamma_{h_2\to l}+\frac{1}{4}\Gamma_{h_1\to h_1}+\frac{1}{4}\Gamma_{h_1\to h_2}+\frac{1}{4}\Gamma_{h_2\to h_1}+\frac{1}{4}\Gamma_{h_2\to h_2}...\right)$$

vs. two separate runs that we bin

$$\frac{1}{2}Tr(...(\not{D}_l^{-1}+\not{D}_{h_1}^{-1})\Gamma(\not{D}_l^{-1}+\not{D}_{h_1}^{-1})...)+\frac{1}{2}Tr(...(\not{D}_l^{-1}+\not{D}_{h_2}^{-1})\Gamma(\not{D}_l^{-1}+\not{D}_{h_2}^{-1})...) = Tr(...$$
$$\Gamma_{l\to l}+\frac{1}{2}\Gamma_{l\to h_1}+\frac{1}{2}\Gamma_{h_1\to l}+\frac{1}{2}\Gamma_{l\to h_2}+\frac{1}{2}\Gamma_{h_2\to l}+\frac{1}{2}\Gamma_{h_1\to h_1}+\frac{1}{2}\Gamma_{h_2\to h_2}...) \tag{2.3}$$

However, this notation is somewhat deceptive since we have $\langle...\Gamma_{h_1\to h_1}...\rangle = \langle...\Gamma_{h_2\to h_1}...\rangle$ since the modes in h_1, h_2 do not have distinct averages. Thus, we obtain

$$\left\langle ...\frac{1}{4}\Gamma_{h_1\to h_1}+\frac{1}{4}\Gamma_{h_1\to h_2}+\frac{1}{4}\Gamma_{h_2\to h_1}+\frac{1}{4}\Gamma_{h_2\to h_2}...\right\rangle = \left\langle ...\frac{1}{2}\Gamma_{h_1\to h_1}+\frac{1}{2}\Gamma_{h_2\to h_2}...\right\rangle$$

justifying our binning procedure in eq. (2.3).

2.0.1.2 Future Work

Since $\not{D}$ empirically doesn't change on another hit (see section 10.6), we could think about doing a block CG (inverting on multiple source vectors). If the memory requirement can be overcome, this would yield a computational advantage.

2.1 Auxiliary Symmetry

We derive a general symmetry between source and sink times lattice diagrams of mesons with degenerate quarks up to four vertices (two-point functions). We apply this symmetry to all-to-all contractions, and show that the number of diagrams computed can be reduced by almost a factor of two. These redundant, or auxiliary diagrams, can then be generated at the analysis stage using only string copying operations (avoiding some round-off errors due to finite precision). While the gains in speed are modest, our main aim in this note is to better understand the underlying combinatorics, given the likelihood of new lattice calculations with large numbers ($\gg 10^3$) of computed diagrams.

2.1.1 Introduction

2.1.1.1 Terminology

We start by defining some terminology. Time separation, or separation, is defined as the separation in time between two source particles. We could also have a sink particle separation, and we set the sink and source separation equal to each other. Thus, we generically refer to separation, or t_{sep}.

Time distance, or distance (t_{dis}), is the distance in time (in the non-looping sense, $(t_{\text{snk}} - t_{\text{src}}) \mod L_t$, L_t is the size of the lattice in the time direction) between the source and sink. If we have non-zero separation, we must pick a particular source time slice and sink time slice to define this distance. We choose the source and sink time slice which are the far apart in time (if time is taken to be positive and we don't consider quark propagators which pass through $t = 0$ to loop around) and refer to these as the inner time slices and the particles on them as inner particles[1] (e.g. inner pion). Clearly this mnemonic is can be confusing because in most cases the particles can loop through the boundary to the other side. However, it will be mathematically well-defined at the beginning of section 2.1.2.1.

Thus, once we have chosen our separation, and we want to compute a correlation function for a particular t_{dis} (regardless of the particles involved), we necessarily have fixed time slices for every source and sink particle. We refer to these locations in time as vertices. In this note, we consider meson two point functions for up to two particles, so examine diagrams of up to four vertices. Each vertex j has a set of quantum numbers associated with it $\{\alpha_i\}_j$. A particular way to connect these vertices with quark propagators is called a topology.

2.1.1.2 Topology Sign Convention/Vertex Layout

If we have vertices $1, 2, ..., n$, we can denote a meson topology by a directed graph where every vertex has one edge coming in and one going out. When we write down the isospin algebra, the coefficient and sign of every distinct topology is fixed. The mapping between these contractions and topologies we call a vertex layout (since much of this occurs at the level of Feynman diagrams).

If all we ever did was sum over equally weighted topologies, we wouldn't need to define a particular layout for our vertices. However, for some correlation functions, e.g. pipi two point with $I = 1$, we have a relative minus sign when between one topology and that topology with source particles fixed and sink particles swapped with each other (or vice-versa). We should fix our vertex layout before we start generating diagrams, but the constructions in this note are independent of vertex layout. This will become clear shortly.

[1] In the particular case of the pipi two point, we refer to the inner pions as pi1, and we thus have a pi1 source and pi1 sink. The outer pions are pi2. This fact is not needed for the purposes of this note.

2.1.1.3 File Structure/Diagram Classes

Consider two particle, isospin symmetric, (meson) two point functions. We choose to separate output from our calculation into files according to conventional definitions of diagram classes $\{R, C, D, V\}$ as in [7]. Thus, R is each particle connected with a source and sink, and no self-contractions. C is each source connected with two sinks. D is each source connected with one sink. V is the disconnected diagram: sources connected to sources, sinks connected to sinks. Within each diagram class, we can have separate topologies. We sum over these topologies (in the g-parity case these are averaged over). In each file, we have matrix elements defined by source time and t_{dis}. Thus, each line in a file has four numbers:

$$t_{\text{src}} \quad t_{\text{dis}} \quad Re(Cor) \quad Im(Cor) \tag{2.4}$$

Importantly, in the naive case, we save separate files for each possible set of $\{\alpha_i\}_j$. The aim of this note is to eliminate this redundancy by nearly a factor of two.

2.1.2 Auxiliary Diagrams

2.1.2.1 Demonstration of Symmetry

Number the vertices and assign quantum numbers:

1. Inner source, $j = 1$, $\{\alpha_i\}_1$, t_{src}
2. Outer source, $j = 2$, $\{\alpha_i\}_2$, $t_{\text{src}} - t_{\text{sep}}$
3. Outer sink, $j = 3$, $\{\alpha_i\}_3$, $t_{\text{src}} + t_{\text{dis}} + t_{\text{sep}}$
4. Inner sink, $j = 4$, $\{\alpha_i\}_4$, $t_{\text{src}} + t_{\text{dis}}$

A particular topology is denoted by, e.g. $1 \to 3 \to 2 \to 4(\to 1)$ (C).

Claim: If we have $\{\alpha_i\}_1 \neq \{\alpha_i\}_3$ or $\{\alpha_i\}_2 \neq \{\alpha_i\}_4$, then we have a single redundancy in our naive diagram case. We can eliminate this auxiliary (aux.) diagram from our lattice calculation. We may need it at the analysis stage, but it can easily be generated from our data set via relatively fast string copy operations. This symmetry holds on a configuration by configuration basis.

To see this, suppose we have without loss of generality $\{\alpha_i\}_1 \neq \{\alpha_i\}_3$. Consider a particular topology for the

R diagram class: $1 \to 3 \to 4 \to 2$. Now, we realize that this topology is also contained in the file that has

$$\{\alpha_i\}_1 \longleftrightarrow \{\alpha_i\}_3 \tag{2.5}$$

$$\{\alpha_i\}_2 \longleftrightarrow \{\alpha_i\}_4 \tag{2.6}$$

However, the corresponding topology is $1 \to 2 \to 4 \to 3$, so we have

$$(1 \to 3 \to 4 \to 2) \longleftrightarrow (1 \to 2 \to 4 \to 3) \tag{2.7}$$

and we must also take

$$t_{\text{src}} \to (t_{\text{src}} + t_{\text{dis}} + t_{\text{sep}}) \mod L_t \tag{2.8}$$

$$\begin{aligned} t_{\text{dis}} &\to (t_{\text{src}} - t_{\text{sep}} - (t_{\text{src}} + t_{\text{dis}} + t_{\text{sep}})) \mod L_t \\ &= (-2t_{\text{sep}} - t_{\text{dis}}) \mod L_t \end{aligned} \tag{2.9}$$

2.1.2.2 Phase Differences

Now, since topologies are summed over, we should be careful so that we don't change relative phases between topologies. Clearly, this is not a problem for two-vertex diagrams. The three vertex diagrams have relative minus signs at the $I = 1$ level, but there are only two such triangle diagram classes (denoted by T). Thus, switching the relative minus sign corresponds to an overall sign flip, so must be worked out once a vertex layout is chosen.

Thus, we are left with the four vertex case. Let's work out the phase change for eq. (2.7) and then generalize to the topologies: Another way to get eq. (2.7) is to reverse the flow, (i.e. take $\to$ and replace with $\leftarrow$). If the vertices are pions, these two topologies have equal weights. In fact, this is true of all pipi topologies (at least as far as the author is aware). Thus, when applying eqs. (2.5) and (2.6) to C, D, V we see D, V topologies are unchanged and C has flow reversed.

The only remaining case is the other two R topologies: $1 \to 4 \to 3 \to 2$ and $1 \to 2 \to 3 \to 4$. They are flow reversed relative to each other. eqs. (2.5) and (2.6) is the identity operation on each of these two topologies. Thus, we see that the particular cases we are examining display this symmetry.

Generally, we find the aux. diagram by applying eqs. (2.5) and (2.6) (usually a filename/directory changing operation) and the following time swaps corresponding to eqs. (2.8) and (2.9) (corresponds to a bijection between lines in the two files).

2.1.2.3 Analysis, Summary

We should really think of this symmetry as a an exact symmetry between source and sink labels. Given a topology, the only freedom we have in labeling it is source vs. sink. This defines a two-fold degeneracy (unless the diagram is symmetric under this interchange). To see this, examine the four vertex diagrams. Note that for realistic L_t, we will have two pairs of vertices that are close in time. If there is an ambiguity in how to pair things up, note that this must mean $t_{\text{dis}} \leq t_{\text{sep}}$. This implies that $L_t <= 4t_{\text{sep}}$, which is unrealistically small for production lattices. Once a given pair is chosen as source, the one that is further along in the positive t direction is necessarily labeled inner, the other outer. At the sink, the one further along in the positive t direction is necessarily labeled outer. As soon as the times labels for the vertices are fixed, the quantum number set index (referred to as j) is fixed as well because the quantum numbers are fixed to a particular vertex before the symmetry operation is applied. This fully fixes the system. If we swap source for sink, we also swap which particle is inner and which is outer, hence eqs. (2.5) and (2.6). This defines the symmetry, up to the physically uninteresting degenerate case where $t_{\text{dis}} = t_{\text{sep}} = \frac{L_t}{4}$. In this case, between two possible pairings we have topologies which swap between diagram classes, making this symmetry impossible to access (at least as the calculations are now set up).

Locating this symmetry in a given lattice calculation is as simple as applying this rule of thumb:

> "If in principle one needs to calculate an amplitude and that same amplitude with source and sink labels swapped, one can save computational time by just calculating one of them."

Then, one must locate the corresponding amplitude ($M(aux)$) in the remainder of the calculation and eliminate it. One can, of course, do this diagrammatically, but it also turns out that (neglecting the symmetry's operation with respect to time) this relationship can be summarized as:

$$M \equiv \phi(p) + ... \rightarrow \psi(p') + ...$$
$$\Rightarrow M(\text{aux}) = \overline{\psi}(-p') + ... \rightarrow \overline{\phi}(-p) + ...$$

which is the definition of crossing symmetry (thanks to A. Meyer for pointing this out initially). As a side note: conjugating the fields ($\phi \Rightarrow \overline{\phi}$) typically doesn't do anything for $\pi\pi$ processes, which may cause one to overlook this similarity initially in this work. The diagrammatic approach to applying the symmetry tends to be less error prone although both approaches are necessarily equivalent. It is unknown at the present time the exact relationship between auxiliary symmetry (a discrete symmetry) and crossing symmetry (a continuum symmetry).

2.1.3 Statistics

What happens when the amplitude is measured on only some (say, every t_{step}) of the source time slices (but all sink time slices)? Clearly, relative to the situation in which both amplitudes are measured, we get less statistics. However, certain time slice measurements will be shared between the measured auxiliary vs. the derived auxiliary amplitude. In our case, we get sharing if (and only if)

$$t_{\text{snk},2} \,\%\, t_{\text{step}} = 0$$

$$t_{\text{snk},2} = t_{\text{src}} + t_{\text{dis}} + t_{\text{sep}}$$

For a fixed $t_{\text{src}}, t_{\text{sep}}$, we get a sharing factor of $\frac{1}{t_{\text{step}}}$. Thus, employing this symmetry, our statistics are reduced by a factor of

$$1 + \left(1 - \frac{1}{t_{\text{step}}}\right)$$

Explanation: In our measured aux correlation function, we get the data from our primary correlation function plus whatever isn't shared by the primary and measured aux, which is one minus the sharing factor. In fact, this factor is universal regardless of the other time slices involved (t_{sep}, etc.) since our sink fields are measured on every time slice by supposition.

As one can see, the statistics hit asymptotically approaches 2 as we measure on fewer and fewer time slices ($t_{\text{step}} \to L_t$). By contrast, if we measure on every time slice, then it makes no statistical difference whether or not we use this crossing symmetry.

2.1.4 G-parity Verification

This symmetry was later used in the G-parity pipi calculation (see (still unpublished at the time of this writing) Lattice 2018 proceedings of C. Kelly). The clearest example of its appearance is in the G-parity D diagram.

2.1.5 Time Swapping Algebra

The time algebra for the symmetry operation is given below in terms of number of vertices and mappings between source times and distances.

2.1.5.1 4 Vertices

(see eqs. (2.8) and (2.9))

2.1.5.2 3 Vertices

$$
\begin{aligned}
t_{\text{src}} &\to (t_{\text{src}} + t_{\text{dis}}) \mod L_t \\
t_{\text{dis}} &\to (t_{\text{src}} - t_{\text{sep}} - (t_{\text{src}} + t_{\text{dis}})) \mod L_t \\
&= (-t_{\text{sep}} - t_{\text{dis}}) \mod L_t
\end{aligned}
$$

2.1.5.3 2 Vertices

$$
\begin{aligned}
t_{\text{src}} &\to (t_{\text{dis}} + t_{\text{src}}) \mod L_t \\
t_{\text{dis}} &\to (t_{\text{src}} - (t_{\text{src}} + t_{\text{dis}})) \mod L_t \\
&= (-t_{\text{dis}}) \mod L_t
\end{aligned}
$$

2.1.6 Concluding Remarks

Why hasn't this symmetry been exploited in previous analyses? Well, take for example the pipi two point function. Let separation be 0. If the pions are stationary, then the quantum numbers are the same at source and sink. Notice also that the above symmetry is independent of t_{src}. In the case of periodic boundary conditions, we see the symmetry we are using is clearly the symmetry between the correlation function at t_{dis} and $L_t - t_{\text{dis}}$, aka the usual periodic symmetry between two halves of an individual file diagram. We thus see duplication of diagrams if we start distinguishing vertices, as in the case of non-stationary mesons, and in cases where the particles are not pions so come with more quantum numbers (as in the rho polarizations).

Additionally, at the level of contractions, speed is usually not an issue. While this symmetry can certainly vastly speed up calculations on limited systems (as on personal computers) with large numbers of diagrams, it is usually a negligible part of the overall computation time. If however the trend of writing a large number of different contractions to disc holds, it may start to compete with other lattice operations (however, right now this seems unlikely).

This symmetry has been checked for periodic boundary conditions, and given that the analysis above is quite general, it could hold for other boundary conditions and particles. The determining factor is clearly the relative phases of the topologies for a given diagram class, which are fixed by physics (as in section 2.1.2.2).

(N.B. This analysis also depends on there being no non-degenerate quarks. In this case, we get more diagram classes corresponding to meson self-contractions.) The rest of the argument is mathematical and so should hold generally.

Net benefits:

1. Since the code to generate auxiliary diagrams from a given data set (with files like eq. (2.4)) parses the last two columns of data as a string, we avoid roundoff errors, which anecdotally, come from how the meson fields are ordered and show up in the last three decimal places of a given correlation function.

2. For some operators, given $\langle A|B^\dagger\rangle$, we need not compute $\langle B|A^\dagger\rangle$.

3. Slightly faster.

2.2 MPI Communication

This work was completed entirely on Intel's Knight's Landing architechture. We refer to the architechture and the supercomputer composed of its nodes as the KNL. Unfortunately, the beginning of our production on this machine was plagued by numerous problems which caused delays. One of the biggest issues with the KNL was the slow inter-node communication speed. Eventually, it was realized that if the communication buffer was composed chiefly of huge memory pages from the Linux kernel, the latency would be reduced resulting in substantial comms performance improvements. (anecdotally a 5x speedup in MPI Broadcast.) We thus undertook to use this buffer in Grid and in our meson field gather code (which was vital considering the large number of broadcasts which occur in a given run).

2.2.1 Loop Hierarchy

We distribute meson fields in a round robin setup. At least one copy of each meson field is kept on a so-called master node. The contraction work is divided up over nodes. We broadcast (in the MPI sense) the fields which are missing on the work nodes. When we are done with a field, we call distribute and it is deleted from all but the master node. However, when we loop over fields, we end up with momentum combinations which are identical save for one difference in one of the particles. In this case, it makes sense to only broadcast the missing field and not call distribute on the remaining (up to 3 for $\pi\pi \to \pi\pi$) meson fields. In practice, arranging the contractions so that we maximize the number of times adjacent momentum combinations only differ by one field is difficult (and such a problem can be shown in general to be NP-complete). We settle for loops over the momentum of source fields and sink fields which are nested so two fields of given momenta are swapped out

at a time. One of the needed meson fields' momentum is derived from a given set of loop indices by momentum conservation (so we only loop over the fields which can be freely chosen given this constraint).

One question one might ask at this point is whether the outer loops should be over sink fields or source fields. We skip (for the connected diagrams) $8-10$ source times to save on compute time. This means most of the work on a given node for a given momentum combination involves a larger variety (in the sense of time) of sink fields than source fields. This means we need to call broadcast many more times for sink fields than source fields. We thus would like to hold the sink fields static and swap only the source fields. To the extent that we can do this and get new work to do, we save on communication time vs. swapping out all the sink fields in our inner loop. Thus, the outer loop(s) should be over sink as opposed to source momenta.

2.3 Split-CG

Generically, we have a set of sources we want to apply $\not{D}^{-1}$ to. If one has a lot of memory but slow comms, it makes sense to parallelize these inversions over nodes. Thus, for each source only a subset of nodes work on the inversion. If the set of sources is said to be fully split, only one node works on a given source. One can invert more sources at once the more nodes one runs on at once. However, the more split the sources are, the more copies of the lattice need to exist in memory, placing limits on the procedure.

2.4 Compressed Eigenvectors

We compress our eigenvectors to save on memory. We save the vectors on a more coarse lattice (and in single precision) as well as the basis change between fine and coarse lattices (although some amount are pure fine grid vectors). This procedure discards short distance information about the eigenvectors. When the eigenvectors are decompressed (translated to coarse grid), we end up with high mode artifacts, but these are quickly suppressed after a few iterations of the solver (since propagators are generally inversely proportional to the energy).

2.5 Data Integrity; Reproduction

We would like to avoid introducing difficult to find numerical bugs in our lattice code. To that end, we must often run reproduction studies, or, at the very least, have evidence that vouches for the integrity of a given ensemble. We should adopt a point of view where we routinely add (hopefully) cheap diagnostics to our code which may not be immediately useful to maximize our chances of not forgetting some crucial aspect of how

some piece of data or analysis was produced.

2.5.1 Git Hash Printing

To these ends, we undertook to encode cryptographic hashes which identify the production means for a given binary. We chose the git hash. Whenever we call make[28], it first writes a header file with a with a C preprocessor definition of the git hash. A minimal compilation unit includes the header file and defines a function which can be called to print the git hash to std::out. We do this for every major library we build from: QMP, Grid, and CPS.

2.5.2 Output File Hashing; Future Work

(As of this writing, the ideas in this section are unimplemented.) Another data integrity problem is how to co-mingle the results from multiple runs that in principle have many of the same input files. In practice, the user might run a binary several times with slightly different input. First, we note that the contents of the dynamic input files are always written to a log. If the user is testing something, they might change some of these input parameters and launch another job in the same directory. Under the current CPS framework, the log of the dynamic input files (so-called 'vml' files) is overwritten. Unless the user takes extra steps to document their work, we may be left with an incomplete history of the data files produced. In the worst case, the user might also finish testing in a given directory and then run a production job. If the output files are not wiped ahead of time, this results in a mix of debug output with final output.

To fix this problem, the dynamic output files should be written to a run-specific log which is never overwritten. As we currently have such a log which documents std::out from a given run, we simply need to add a function which pipes the contents of the vml files to std::out.

Another data integrity problem is tying the log files to the output data files. To solve this, we could hash the output files and print the hashes in the output log. We then could, at the end of the program, hash the output log and print that hash in one of the last lines. At that point, any alternation of the log or the data files will yield an inconsistency in the records (allowing us to mark a dataset as dirty).

2.6 Sample AMA

A procedure of C.Kelly (private comm.) named Sample AMA allows us to extend the well-known AMA procedure[29] to entire sets of measurements performed on a given gauge configuration. The usual AMA procedure

is to form an error corrected quantity

$$\langle O\rangle_{\text{ imp}} = \langle O\rangle_{\text{ exact}} - \langle O\rangle_{\text{ approx}} + \frac{1}{N}\sum_{i=1}^{N}\langle O\rangle_{i,\text{ approx}}$$

where the $O_{\text{ approx}}$ have some systematic error but have much greater statistics (a factor of N more) compared to $\langle O\rangle_{\text{ exact}}$. We thus trade systematic error for statistical. However, it is usually much cheaper in terms of computational cost to calculate $O_{\text{ approx}}$ (also referred to as the sloppy measurements), so we are able to get much better statistics for equal computational cost of only computing $O_{\text{ exact}}$. AMA is usually defined by finding greater statistics of $\langle O\rangle_{\text{ approx}}$ via some lattice symmetry like time translation symmetry. Thus, we might measure on more $t_{\text{ src}}$ for our approximate measurement (and i would refer to a $t_{\text{ src}}$). Sample AMA simply extends this notion to gauge symmetry.

We form a jackknifed quantity $\overline{A}_i$ with standard error denoted as σ the usual way:

$$\overline{A}_i = \frac{1}{N_A - 1}(N_A \langle A\rangle - A_i) \tag{2.10}$$

We then can form an expanded quantity:

$$\tilde{A}_i = \begin{cases} \overline{A}_i & i < N_1 \\ \langle A\rangle & N_1 <= i < N_1 + N_2 \end{cases}$$

Clearly, $\langle A\rangle = \langle \tilde{A}\rangle$. The jackknife standard error is

$$\begin{aligned}\tilde{\sigma}_A^2 &= \frac{N_1+N_2-1}{N_1+N_2}\sum_{i=0}^{N_1+N_2-1}\left(\tilde{A}_i - \langle A\rangle\right)^2 \\ &= \frac{N_1+N_2-1}{N_1+N_2}\sum_{i=0}^{N_1-1}\left(\overline{A}_i - \langle A\rangle\right)^2 \\ &= \frac{N_1+N_2-1}{N_1+N_2}\frac{N_1}{N_1-1}\sigma^2\end{aligned}$$

We can do the same thing for another quantity:

$$\tilde{B}_i = \begin{cases} \langle B \rangle & i < N_1 \\ \overline{B}_i & N_1 <= i < N_1 + N_2 \end{cases}$$

If our result is some quantity $A+B$, we can see the error on the result comes from adding the errors on the extensions in quadrature:

$$\tilde{\sigma}^2 = \tilde{\sigma}_A^2 + \tilde{\sigma}_B^2 \tag{2.11}$$

Now, the procedure is to do exact and sloppy measurements on a subset of the ensemble of amount N_1. We then form the correction by subtracting the sloppy from the exact. We therefore identify A,B with this correction and sloppy only measurements respectively. The resulting set we refer to as the superjackknife blocks. These clearly have standard error eq. (2.11).

2.7 Even-Odd Preconditioning

Using the fact that the Dirac operator is a highly sparse matrix, it turns out we can write a preconditioned Dirac operator that separates the inversion problem into parts that are trivial to invert (because they are essentially diagonal) and parts which are non-trivial. This even-odd preconditioning reduces the size of our problem by roughly a factor of 2. We can see this from looking at the non-trivial part $(D_{oo}^\dagger D_{oo})^{-1} - \sum \frac{1}{\lambda_i} h_i h_i^\dagger$ in eq. (2.12). We summarize some necessary modifications we made to the code for the Dirac operator's even-odd preconditioning. These modifications took place in the course of a general period of time when most of the CPS library was being rewritten for SIMD as Grid. Incompatibilities necessitated code changes like the one detailed in this section.

We write down how to change from Asym preconditioning to Sym2 preconditioning for v,w vectors of the all-to-all quark propagator.

2.7.1 Definitions

The even-odd preconditioned Dirac operator can be defined several ways. We define two. The notation is similar to p. 60 of [30].

2.7.1.1 Asym

[30] uses Asym preconditioning, which we restate here

$$
\begin{aligned}
D_{DWF} &= \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \begin{pmatrix} M_{ee} & 0 \\ 0 & D_{oo}^{\text{Asym}} \end{pmatrix} \begin{pmatrix} 1 & M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \\
D_{oo}^{Asym} &\equiv M_{oo} - M_{oe}M_{ee}^{-1}M_{eo}
\end{aligned}
$$

The inverted Dirac operator is then

$$
D_{DWF}^{-1} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_{ee}^{-1} & 0 \\ 0 & (D_{oo}^{\dagger\,\text{Asym}} D_{oo}^{\text{Asym}})^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & D_{oo}^{\dagger\,\text{Asym}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix}
$$

2.7.1.2 Sym2

$$
\begin{aligned}
D_{DWF} &= \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \begin{pmatrix} M_{ee} & 0 \\ 0 & D_{oo}^{\text{Sym2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & M_{oo} \end{pmatrix} \begin{pmatrix} 1 & M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \\
D_{oo}^{\text{Sym2}} &\equiv 1 - M_{oe}M_{ee}^{-1}M_{eo}M_{oo}^{-1}
\end{aligned}
$$

which causes the inverted operator to be decomposed as

$$
D_{DWF}^{-1} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & M_{oo}^{-1} \end{pmatrix} \begin{pmatrix} M_{ee}^{-1} & 0 \\ 0 & (D_{oo}^{\dagger\,\text{Sym2}} D_{oo}^{\text{Sym2}})^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & D_{oo}^{\dagger\,\text{Sym2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix}
$$

2.7.2 Transformation Factors

We thus see that the transformation factor which must be applied to go from Asym to Sym2 is a factor multiplying the A2A vector v_i. This factor is the correction matrix (highlighted in section 2.7.1). We thus can apply the following transformation to the expression for v_i given in [30].

2.7.2.1 Low Modes

$$v_i^{\text{Asym}} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\lambda_i} h_i \end{pmatrix}$$

$$v_i^{\text{Sym2}} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & M_{oo}^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\lambda_i} h_i \end{pmatrix}$$

2.7.2.2 High Modes

The Möbius-only operator D_- is included. If not using Möbius, ignore this.

$$v_i^{\text{Asym}} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_{ee}^{-1} & 0 \\ 0 & (D_{oo}^\dagger D_{oo})^{-1} - \sum \frac{1}{\lambda_i} h_i h_i^\dagger \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & D_{oo}^\dagger \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} (D_-)\eta_i$$

$$v_i^{\text{Sym2}} = \begin{pmatrix} 1 & -M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & M_{oo}^{-1} \end{pmatrix} \begin{pmatrix} M_{ee}^{-1} & 0 \\ 0 & (D_{oo}^\dagger D_{oo})^{-1} - \sum \frac{1}{\lambda_i} h_i h_i^\dagger \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & D_{oo}^\dagger \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} (D_-)\eta_i \tag{2.12}$$

2.7.3 Grid Functions

We here describe how to use Grid's[31] code conventions to implement preconditioning. To implement this in Grid, note that each linear operation Op is defined as

$$Op(A,B) \equiv$$

$$B := Op \cdot A$$

There are two general operators types for the M matrices given above. One is independent of precondition scheme:

$$Meooe = M_{eo}, M_{oe}$$

$$Mooee = M_{oo}, M_{ee}$$

$$MooeeInv = M_{oo}^{-1}, M_{ee}^{-1}$$

$$MeooeInv = M_{eo}^{-1}, M_{oe}^{-1}$$

Which operation in the pair they perform depends on the grid selection (odd or even) made for the vector (hence the commas).

The other depends on the precondition scheme (equivalently, the A2A policy defined at compile time):

$$D \equiv Mpc$$

$$D^\dagger \equiv MpcDag$$

$$D^\dagger D \equiv HermOp$$

Since the current A2A library in CPS mixes precondition independent and precondition dependent operators, we must apply a correction factor to the precondition independent operators when switching schemes. In this case, the correction factor is needed to go from Asym to Sym2.

We thus can see that the correction factor for v_i's that should be applied second to last is

$$MooeeInv$$

Chapter 3

Ensemble Details

3.1 16^3, Consistency Comparison Ensemble

Table 3.1: A2A parameters

Parameter	[30]	Updated+Grid
nl	100	–
nhits	1	–
src width	1	–
t sep	4	–
pion radius	2.0	–
L_t	32	–
beta	2.13	–
Ls	16	–
Ms	0.032	–
Mu	0.01	–
# Configurations	400	643
Nh	?	768
Random num type	?	U(1)
M_5	–[1]	1.80
BfmSolverType solver	?	BFM_HmCayleyTanh (DWF)
möbius scale	?	1.0
Gauge	?	Iwasaki
Gauge Fix	?	Coulomb
Gauge Fix rsd	?	1e-08
Max its, GF	?	20000
Lanczos rsd	?	1e-08
Lanczos, Max its	?	10
CG stop rsd	?	1e-16

We use a 16^3 ensemble to check our contraction code with measurements performed by Daiqan Zhang[32]. Parameters common to both of our ensembles are listed first. Values which are unknown to us are listed

as question marks. Values which are the same in our ensemble are denoted with –.

Explanation of some of the parameters (the rest are explained in section 3.2):

nhits is the number of random stochastic source vectors used per mode. We use only one, as the sum over configurations is also used to perform this average.

Source width the width in time of the source (obsolete – always set to one now, source isolated to single time slice).

t sep is the separation in time of pions. This method was pioneered by Q. Liu([7]) to reduce vacuum contributions to the correlation function. Two particles (e.g. two pions) are forced to not annihilate or create on the same time slice, but be separated at source and sink by t sep.

3.2 24^3 DSDR

The following are the production parameters for the coarser lattice. This lattice is cheaper to calculate on than section 3.3. The Dislocation Suppressing Determinant Ratio (DSDR) lattice has a force term which suppresses chiral symmetry breaking terms that are enhanced at strong coupling and very coarse lattices. See [33],[34] for further information.

Table 3.2: Production Parameters

Parameter	Value
Lattice size	24^3 *x* 64 *x* 12
coarse grid block size	12 *x* 2 *x* 3 *x* 3 *x* 2
Spatial Boundary Conditions	Periodic
Time Boundary Conditions	Anti-periodic
nl	2000
nl, fine grid	1000
a^{-1}	1.015 GeV
tdis max	16
π, σ, ρ radius	1.5
L_t	64
beta	1.633
Ls (zMöbius approx of Möbius Ls=24, Shamir Ls=96)	12
Ms	0.0850
Mu	1.07e-03
Tstep	8
Tsep (see section C.3 for definition)	3
# Configurations (sloppy only, no MADWF)	178
Nh	768
Random num type	U(1)
M_5	1.80
Gauge	Iwasaki+DSDR
Gauge Fix	Coulomb
Gauge Fix residual	1e-14 (see eq. 8 in [9])
Lanczos rsd	1e-06
CG stop rsd	1e-08
CG method (sloppy, zmöbius)	Single-prec, 400 iterations
Avg. CG True Resid (sloppy)	1e-06
RNG Seed	(fixed to config number)

Production parameters for the 24^3 lattice ensemble.

Explanation of parameters:

Lattice size : in lattice units we have $V^3 \times L_t \times L_s$

coarse grid block size We compress eigenvectors by projecting to a coarser lattice. This is the size of the coarse lattice.

nl is the number of eigenvectors we use (to deflate our Dirac operator). See section 2.0.1. nl fine grid is the number of eigenvectors we compute on the fine grid. The rest are computed on the coarse grid.

a^{-1} is the inverse lattice spacing

$t_{\textbf{dis}}$ is the time distance between inner pions (pions closest together in time when travelling forward in time from source to sink). Tdis max is the maximum t_{dis} we measure. Beyond this point we assume the data is too noisy to be useful.

radius We use a hydrogen-like wavefunction with spatial dependence $exp(-x/r)$ where x is the relative distance between quark and anti-quark in our meson and r is the radius.

beta is a term associated with the Iwasaki gauge action

Ls is the extent of the fifth dimension associated with using domain wall fermions.

Ms is the strange quark mass in lattice units.

Mu is the up and down quark mass in lattice units.

Tstep is the number source time slices we skip when placing our source operator. Because of time translation symmetry, we should get the same answer for our correlation functions regardless of the source time slice we place our source operator.

Tsep see section 3.1

MADWF We must AMA correct (see section 2.6) our zMöbius action using the MADWF (see section 1.3.2) procedure on a subset of our 185 gauge configurations.

Nh is the number of spatially distinct random source vectors we use in section 2.0.1. We refer to it as the number of high modes since these usually have small overlap with the low mode spectrum of the Dirac operator.

U(1) means we use a random ϕ in $e^{i\phi}$ for every spatial site in a given high mode source.

M_5 is the domain wall fermion mass (or, roughly speaking, the height of the domain wall)

Gauge fix Coulomb means we get rid of our redundant gauge degree of freedom by enforcing $\partial_i A_i^a = 0$.

Gauge fixing residual is the stopping condition used.

Lanczos residual is the norm dependent (for operator A with eigenvector, eigenvalue x, λ: $|Ax - \lambda x|$) residual for our low modes calculated via the well-known Lanczos algorithm.

CG stop rsd is the same quantity as the Lanczos but for the solver and calculated in a norm independent way: $|Ax - b|/|x|$

Single prec, 400 iterations For our sloppy action (non-MADWF) we need not perform as many iterations since it is useless past the accuracy of the operator. We would like to avoid large jumps in residual around the cutoff point, so we watch a few inversions on different configurations and then choose a conjugate gradient iteration number as our stopping point (assuming the residual is always stable in this region).

RNG Seed Our A2A noise vectors (see section 2.0.1) are generated using a random number generator. We need a fixed integer seed for reproducibility, so for simplicity the author chose to fix it to the gauge configuration number. Additional A2A hits on a single gauge configuration are generated by incrementing this integer.

3.3 32ID-Fine

The following are the production parameters for the finer lattice we use to take the continuum limit.

Table 3.3: Production Parameters

Parameter	Value
Lattice size	32^3 *x* 64 *x* 12
coarse grid block size	12 *x* 2 *x* 2 *x* 2 *x* 2
Spatial Boundary Conditions	Periodic
Time Boundary Conditions	Anti-periodic
nl	2000
nl, fine grid	250
a^{-1}	1.015 GeV
tdis max	16
π, σ, ρ radius	1.5
L_t	64
beta	1.75
Ls (Shamir Ls=32)	12
Ms	0.0850
Mu	0.0001
Tstep	10
Tsep (see section C.3 for definition)	3
Tdis max	22
π radius	2.0
σ radius	2.0
ρ radius	2.0
# Configurations (sloppy only, no MADWF)	66
Nh	768
Random num type	U(1)
M_5	1.80
Gauge	Iwasaki+DSDR
Gauge Fix	Coulomb
Gauge Fix residual	1e-14 (see eq. 8 in [9])
Lanczos rsd	1e-06
CG stop rsd	1e-08
Avg. CG True Resid (sloppy)	1e-05
CG method (sloppy, möbius)	Single-prec, 330 iterations
CG method (exact)	Single-prec, 1400 iterations single, with restarted double precision
Möbius $b+c$	$\frac{8}{3}$
RNG Seed	(fixed to config number)

Production parameters for the 32^3 lattice ensemble.

Additional parameter unique to the 32^3:

Mobius $b+c$ These are the analogue of the zMöbius parameters (see section 1.3.2 and chapter G), except these are constant in s.

Chapter 4

$\pi\pi, \sigma$ Contractions

4.1 Clebsch-Gordan Coefficients

We are interested in $I = 0$ and $I = 2$ from the $1x1$ blocks, but we want to reduce the representations to the quark model states. Since $\begin{pmatrix} u \\ d \end{pmatrix}$ must transform under an isospin rotation U the same as $\begin{pmatrix} \overline{u} \\ \overline{d} \end{pmatrix}$ under U^*, we would like to have an antiquark column vector that transform like $\begin{pmatrix} u \\ d \end{pmatrix}$. It turns out that this is $\begin{pmatrix} \overline{d} \\ -\overline{u} \end{pmatrix}$. Thus, we expand the isospin states down to the quark level, and then we replace a $\uparrow$ (see, e.g., section 4.1.1 for use of $\uparrow, \downarrow$ notation referring to up and down spin in a typical $SU(2)$ CG decomposition which we are mapping to isospin through the use of this notation) with a $-\overline{d}$. For a $\overline{u}$, we replace a $\downarrow$, and drop an overall phase constant. (To see this, write out the expansion for a general rotation in two dimensional SU(2). Take the complex conjugate of both sides. We want to find a matrix such that $M\overline{\psi}$ transforms under U like $\overline{\psi}$ under U^*. The defining equation is thus $MU^* = UM \Rightarrow M = phase * \sigma^2$)

$$\psi' = U\psi$$

$$\Rightarrow \overline{\psi}' = U^*\overline{\psi}$$

$$M\overline{\psi}' = UM\overline{\psi}$$

$$\Rightarrow MU^* = UM$$

4.1.1 I=0

$$
\begin{aligned}
|0,0\rangle &= \frac{1}{\sqrt{3}}(|1,1\rangle|1,-1\rangle - |1,0\rangle|1,0\rangle + |1,-1\rangle|1,1\rangle) \\
&= \frac{1}{\sqrt{3}}\left(\pi^+\pi^- - \pi^0\pi^0 + \pi^-\pi^+\right) \\
&= \frac{1}{\sqrt{3}}\left(\uparrow\uparrow\downarrow\downarrow - \frac{1}{2}(\uparrow\downarrow\uparrow\downarrow + \uparrow\downarrow\downarrow\uparrow + \downarrow\uparrow\uparrow\downarrow + \downarrow\uparrow\downarrow\uparrow) + \downarrow\downarrow\uparrow\uparrow\right) \\
&= \frac{1}{\sqrt{3}}\left(u\bar{d}d\bar{u} + \frac{1}{2}\left(u\bar{u}u\bar{u} - u\bar{u}d\bar{d} - d\bar{d}u\bar{u} + d\bar{d}d\bar{d}\right) + d\bar{u}u\bar{d}\right)
\end{aligned}
$$

4.1.2 I=2

$$
\begin{aligned}
|2,2\rangle &= |1,1\rangle|1,1\rangle \\
&= \pi^+\pi^+ \\
&= \left|u\bar{d}\right\rangle\left|u\bar{d}\right\rangle \\
|2,1\rangle &= \frac{1}{\sqrt{2}}(|1,1\rangle|1,0\rangle + |1,0\rangle|1,1\rangle) \\
&= \frac{1}{\sqrt{2}}\left(\pi^+\pi^0 + \pi^0\pi^+\right) \\
&= \frac{1}{2}(\uparrow\uparrow\uparrow\downarrow + \uparrow\uparrow\downarrow\uparrow + \uparrow\downarrow\uparrow\uparrow + \downarrow\uparrow\uparrow\uparrow) \\
&= \frac{1}{2}\left(-u\bar{d}u\bar{u} + u\bar{d}d\bar{d} - u\bar{u}u\bar{d} + d\bar{d}u\bar{d}\right) \\
|2,0\rangle &= \frac{1}{\sqrt{6}}(|1,-1\rangle|1,1\rangle + 2|1,0\rangle|1,0\rangle + |1,1\rangle|1,-1\rangle) \\
&= \frac{1}{\sqrt{6}}\left(\pi^-\pi^+ + 2\pi^0\pi^0 + \pi^+\pi^-\right) \\
&= \frac{1}{\sqrt{6}}(\downarrow\downarrow\uparrow\uparrow + \uparrow\downarrow\uparrow\downarrow + \uparrow\downarrow\downarrow\uparrow + \downarrow\uparrow\uparrow\downarrow + \downarrow\uparrow\downarrow\uparrow + \uparrow\uparrow\downarrow\downarrow) \\
&= \left(-d\bar{u}u\bar{d} + u\bar{u}u\bar{u} - u\bar{u}d\bar{d} - d\bar{d}u\bar{u} + d\bar{d}d\bar{d} - u\bar{d}d\bar{u}\right) \\
|2,-1\rangle &= \frac{1}{\sqrt{2}}\left(\pi^-\pi^0 + \pi^0\pi^-\right) \\
&= \frac{1}{\sqrt{2}}\left(u\bar{u}d\bar{u} - d\bar{d}d\bar{u} + d\bar{u}u\bar{u} - d\bar{u}d\bar{d}\right) \\
|2,2\rangle &= \pi^-\pi^- \\
&= d\bar{u}d\bar{u}
\end{aligned}
$$

4.2 Contractions, Diagrams, Types

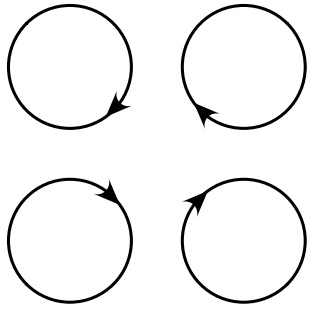
4.2.1 Irrelevant diagrams

The diagrams of sec. 4.2.1.1, 4.2.1.2, and 4.2.1.3 all cancel out. proof: Each of these diagrams has a pion contracted with itself. This pion must be a π^0. This means that this pion must either be flavored $d\overline{d}$ or flavored $u\overline{u}$. Either way, there must exist the same diagram with the other flavor and a relative minus sign. Since we are working in the degenerate mass limit, these two diagrams cancel out.

We simplify the algebra by first listing the diagrams with a sketch of their respective traces and assigning a variable to each. We also work with degenerate quark masses.

4.2.1.1 4 Bubbles

Or at most 1 contracted.

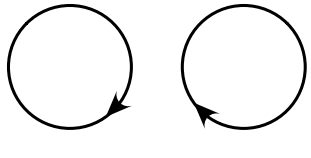


$$= Tr\left(\gamma_5 M^{-1}\right)^4$$

$$\equiv a$$

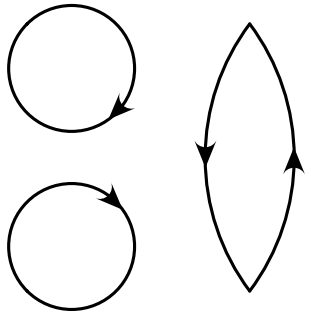
4.2.1.2 3 Bubbles

Or at most 2 contracted.



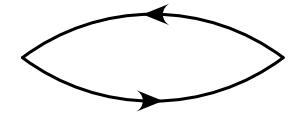
$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{12}$$



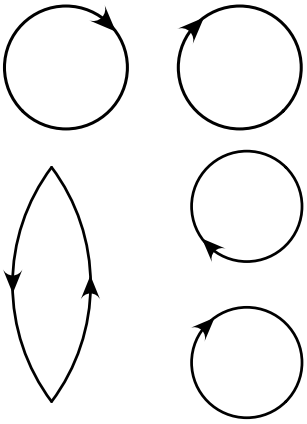
$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{23}$$



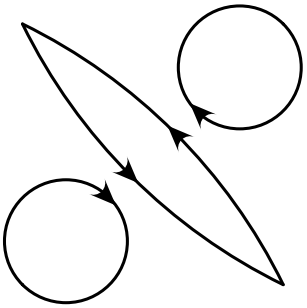
$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{34}$$



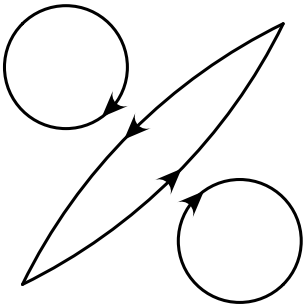
$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{41}$$



$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{13}$$

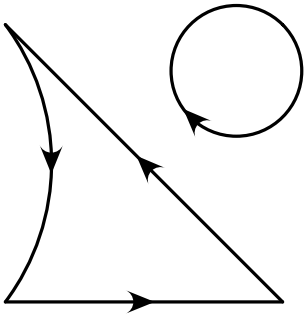


$$= Tr\left(\gamma_5 M^{-1}\right)^2 Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv b_{24}$$

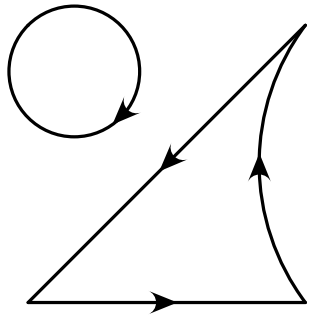
4.2.1.3 1 Bubble

Or at most 3 contracted



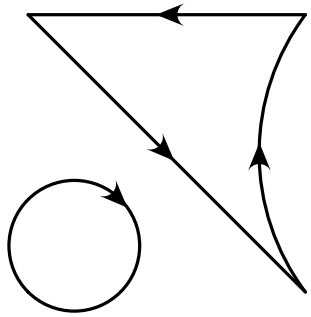
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_1$$



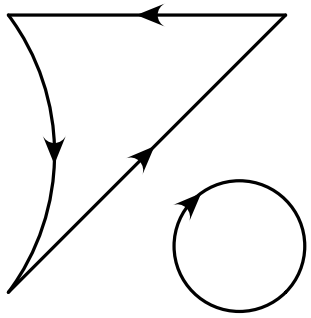
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_2$$



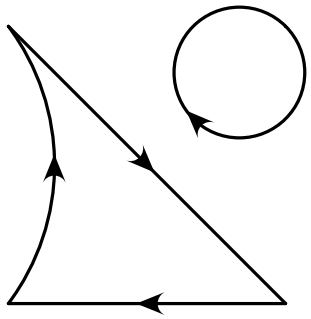
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_3$$



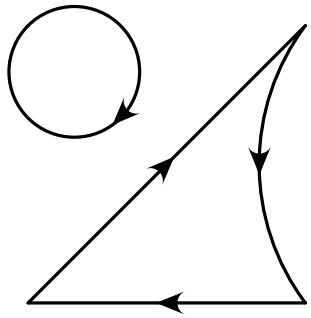
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_4$$



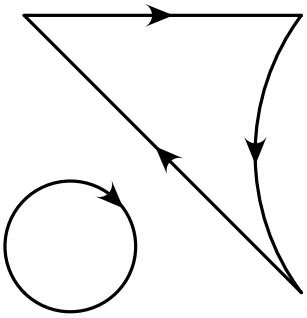
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_{1r}$$



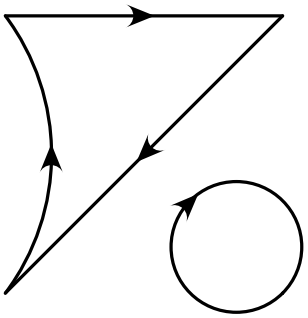
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_{2r}$$



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

$$\equiv d_{3r}$$



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right) Tr(\gamma_5 M^{-1})$$

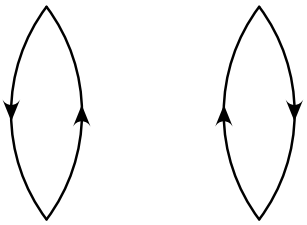
$$\equiv d_{4r}$$

We have d_{1r}, d_{2r}, d_{3r} and d_{4r} , which correspond to d_1, d_2, and d_3 respectively, with particle flow reversed (clockwise to counterclockwise, counterclockwise to clockwise). These are distinct contractions, and need to be listed. Notice that in previous sections, particle flow could be reversed, but we get the same contraction back (a through c).

4.2.2 Relevant Diagrams

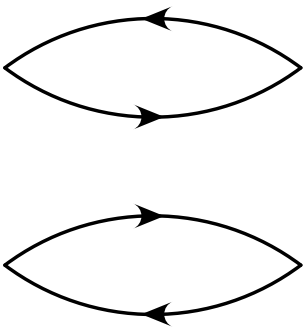
4.2.2.1 2 Bubbles

Or at most 2 contracted (rest).



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)^2$$

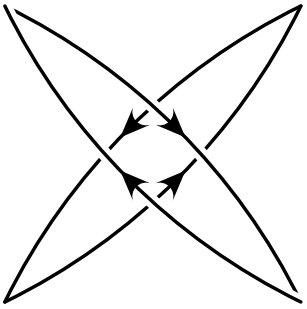
$$\equiv c_v$$



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)^2$$

$$\equiv c_h \qquad (4.1)$$

(4.2) below must have the same contribution to the brackets we are calculating since it is simply (4.1) with outgoing pions swapped.

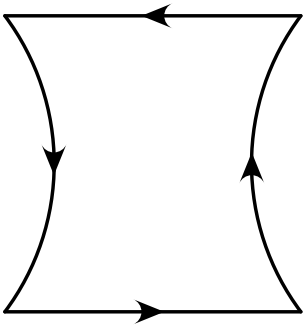


$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\right)^2$$

$$\equiv c_c \tag{4.2}$$

4.2.2.2 Squares

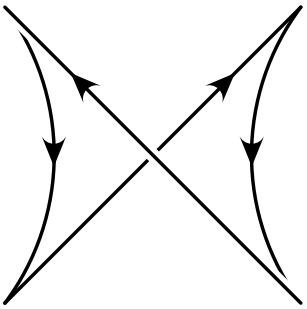
Or at most 4 contracted.



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

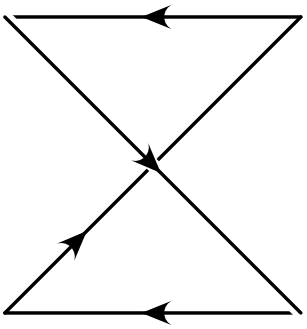
$$\equiv e_1 \tag{4.3}$$

(4.3) must contribute to the calculated brackets the same as (4.4) since it simply has outgoing particles swapped. (same reasoning as the (4.1) and (4.2) equivalence)



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv e_2 \tag{4.4}$$

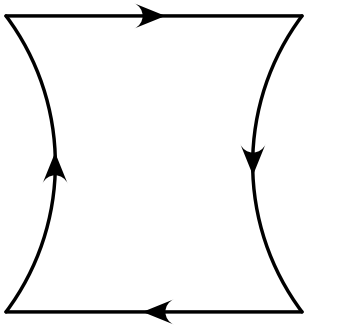


$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv e_3$$

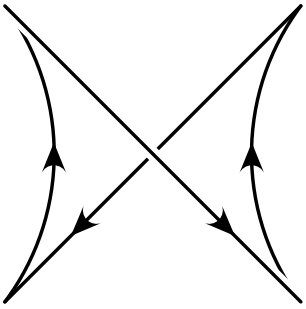
This last section's connections (contractions without direction) are a little bit tricky to see, but if one writes the numbers all permutations of the numbers $1 \to 4$, we see three distinct orderings up to cyclic re-orderings.

4.2.3 Reversed[1], Relevant Diagrams



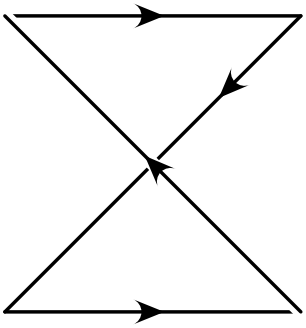
$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv e_{1r}$$



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv e_{2r}$$



$$= Tr\left(\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\right)$$

$$\equiv e_{3r}$$

We have, for example, e_{1r}, e_{2r}, and e_{3r} , which correspond to e_1, e_2, and e_3 with particle flow reversed. The diagrams have the same amplitude as their non-reversed counterparts, but we need to include them to make sure we end up with the proper number of contractions:

We end up with 24(= 4!) total diagrams if we add up each separate diagram from sections 4.2.1 and 4.2.2, as expected. (If all quarks are the same flavor, we have 4 choices for the first contraction, 3 for the second and so on.)

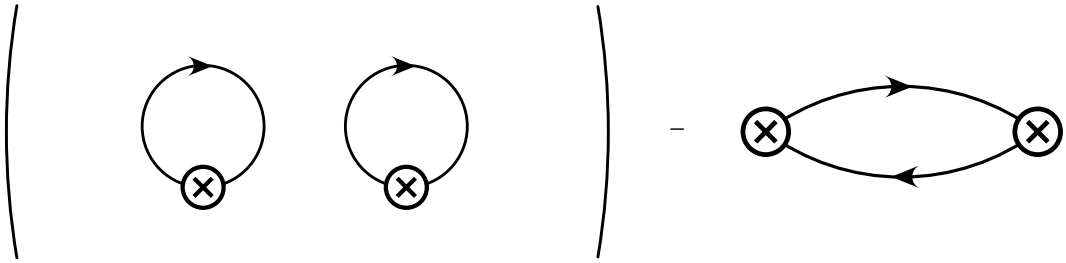
4.3 $I = 0$ Contractions

4.3.1 Scalar-Scalar

Contractions for the scalar to scalar vev (where we know the $\sigma = \frac{1}{\sqrt{2}}\left(\overline{u}u + \overline{d}d\right)$. The normalization of $\sigma \to \sigma$ is $\frac{1}{2}$ and we get 4 separate bubble contractions and 2 separate connected contractions. The Grassmann algebra

[1] Reversed means in the sense of particle flow arrows

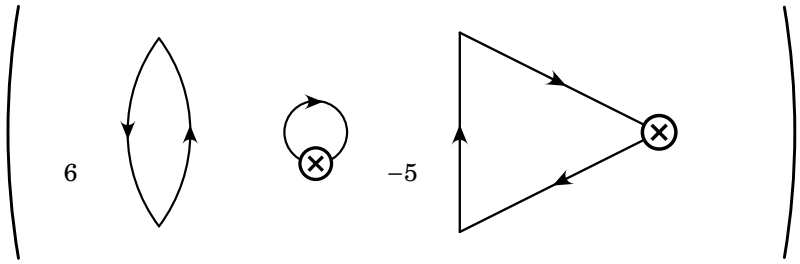
give us the minus sign on the connected diagram).

$\langle\sigma|\sigma\rangle =$ 2  −

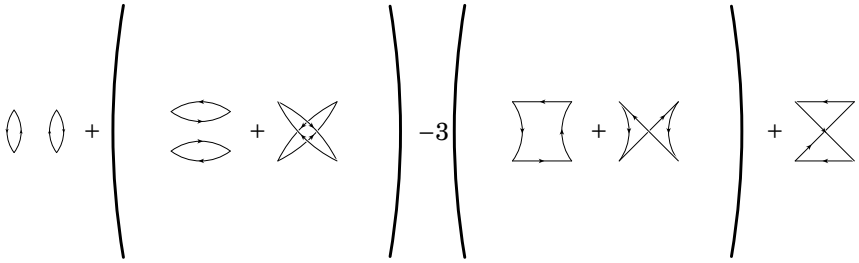
4.3.2 Pion-Scalar

Scalars are denoted by a circle with an "x" through it.

Contractions for pion to scalar vev.

$\langle\sigma|\pi, I=0\rangle = \frac{1}{\sqrt{6}}$ 

4.3.3 Pion-Pion

$\langle\pi, I=0|\pi, I=0\rangle =$ 3 

4.3.3.1 Clebsch-Gordan Bracket Algebra

We show the CG algebra for $(\pi\pi)_{I=0} \to (\pi\pi)_{I=0}$.

$$\langle 0|0\rangle = \frac{1}{3}(\langle 1,1;1,-1|1,1;1,-1\rangle + \langle 1,1;1,-1|1,-1;1,1\rangle + \langle 1,-1;1,1|1,-1;1,1\rangle + \langle 1,-1;1,1|1,1;1,-1\rangle)$$
$$+ \frac{1}{3}(\langle 1,0;1,0|1,0;1,0\rangle - \langle 1,1;1,-1|1,0;1,0\rangle - \langle 1,-1;1,1|1,0;1,0\rangle - \langle 1,0;1,0|1,1;1,-1\rangle - \langle 1,0;1,0|1,-1;1,1\rangle)$$
$$= \frac{1}{3}\langle 0|\left(\overline{d}\gamma_5 u\overline{u}\gamma_5 d + \frac{1}{2}\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right) + \overline{u}\gamma_5 d\overline{d}\gamma_5 u\right)$$
$$\left(\overline{u}\gamma_5 d\overline{d}\gamma_5 u + \frac{1}{2}\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right) + \overline{d}\gamma_5 u\overline{u}\gamma_5 d\right)|0\rangle$$

(We drop the overall $\frac{1}{3}$ until the end.)

$$= \langle|\,\overline{d}\gamma_5 u\overline{u}\gamma_5 d\overline{u}\gamma_5 d\overline{d}\gamma_5 u\,|\rangle + \langle|\,\overline{d}\gamma_5 u\overline{u}\gamma_5 d\overline{d}\gamma_5 u\overline{u}\gamma_5 d\,|\rangle + \langle|\,\overline{u}\gamma_5 d\overline{d}\gamma_5 u\overline{u}\gamma_5 d\overline{d}\gamma_5 u\,|\rangle + \langle|\,\overline{u}\gamma_5 d\overline{d}\gamma_5 u\overline{d}\gamma_5 u\overline{u}\gamma_5 d\,|\rangle \tag{4.5}$$
$$+ \frac{1}{2}\langle|\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)\overline{u}\gamma_5 d\overline{d}\gamma_5 u\,|\rangle \tag{4.6}$$
$$+ \frac{1}{2}\langle|\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)\overline{d}\gamma_5 u\overline{u}\gamma_5 d\,|\rangle \tag{4.7}$$
$$+ \frac{1}{2}\langle|\,\overline{d}\gamma_5 u\overline{u}\gamma_5 d\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)|\rangle \tag{4.8}$$
$$+ \frac{1}{2}\langle|\,\overline{u}\gamma_5 d\overline{d}\gamma_5 u\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)|\rangle \tag{4.9}$$
$$+ \frac{1}{4}\langle|\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)\left(\overline{u}\gamma_5 u\overline{u}\gamma_5 u - \overline{u}\gamma_5 u\overline{d}\gamma_5 d - \overline{d}\gamma_5 d\overline{u}\gamma_5 u + \overline{d}\gamma_5 d\overline{d}\gamma_5 d\right)|\rangle \tag{4.10}$$

Shortcuts:

We work in the limit of degenerate quark masses, so we pick up a u,d symmetry factor of 2 when we add a term to the same term with $u \to d$ and $d \to u$. Also, it can be shown that the sign of each contraction never changes. Thus, all the e's come with a -1 and all the c's come with a $+1$. To read off the contractions from a vacuum expectation value of 4 pion operators, look at the list of relevant diagrams and do the following

1. The numbering convention in this document is by vertex. See section 2.1.2.1 for the convention. Otherwise, pick a convention.
2. The corresponding pions in the brackets are $4 \to 1$ from left to right.
3. Choose the least numerous quark-antiquark pair (u or d).
4. Do all of the contractions as follows:

(a) If there are no pairs of this quark, then we are done: If it is all ups or all downs all you have to do is read off all the relevant diagrams (see section (section 4.2.2)).

(b) if there is only one pair of this quark, do the contraction of this pair. Whichever two positions the pair occupies, read off the relevant diagrams with this line.

(c) if there are an equal number of ups and downs, then we proceed to do the contractions for the downs only. these two contractions yield two contractions which can be picked out from the list of relevant diagrams. A convention regarding particle flow is generally helpful, but for $\pi\pi \to \pi\pi$ it turns out not to matter.

Contractions for (4.5):

$$\begin{aligned}
&\pi^+\pi^- \to \pi^+\pi^- + permutations \\
&= 2(c_v - e_{1r} - e_1 + c_h) + 2(c_v + c_c - e_2 - e_{2r}) \\
&= (4c_v + 2c_c + 2c_h) + (-2e_1 - 2e_2 - 2e_{1r} - 2e_{2r})
\end{aligned} \tag{4.11}$$

Permutations indicates that we are summing over the four possible versions of this process (exchanging the two final state pions with each other, or exchanging the two initial state pions with each other.)

Contractions for (4.6):

$$\begin{aligned}
&\pi^-\pi^+ \to \pi^0\pi^0 \\
&= (-b_{41} + d_{4r} + c_v - e_{1r} - e_{2r} + d_{1r}) + (-b_{41} + d_4 + c_v - e_2 - e_1 + d_1) - (-b_{41} + d_4 + d_{1r} - e_3) - (-b_{41} + d_{4r} + d_1 - e_{3r}) \\
&= (2c_v) + (-e_1 - e_2 - e_{1r} - e_{2r} + e_3 + e_{3r})
\end{aligned} \tag{4.12}$$

The $\frac{1}{2}$ cancels with the u,d symmetry factor, 2.

Contractions for (4.7):

$$\pi^+\pi^- \to \pi^0\pi^0$$

Already accounted for. (These give the same amplitude as (4.6), so are the reason the $\frac{1}{2}$ cancels in section 4.12.)

Contractions for (4.8):

$$\begin{aligned}&\pi^0\pi^0 \to \pi^-\pi^+\\&= (2c_v) + (-e_1 - e_2 - e_{1r} - e_{2r} + e_3 + e_{3r})\end{aligned} \tag{4.13}$$

The $\frac{1}{2}$ cancels with the u,d symmetry factor, 2. We swap initial and final states with (4.6), which corresponds to reflecting the corresponding diagrams across the vertical axis. Since the diagram sum of (4.6) is already symmetric across the vertical axis, we get the same contractions as (4.6):

Contractions for (4.9):

$$\pi^0\pi^0 \to \pi^+\pi^-$$

Already accounted for. (These give the same amplitude as (4.8), so are the reason the $\frac{1}{2}$ cancels in section 4.13.)

Contractions for (4.10):

$$\pi^0\pi^0 \to \pi^0\pi^0$$

We list contractions by the number of the term from the first parentheses and the number of the term from the second parentheses. For example, "24" is

$$-\frac{1}{4}\langle|\, u\gamma_5\overline{u}d\gamma_5\overline{d}d\gamma_5\overline{d}d\gamma_5\overline{d}\,|\rangle$$

11+44:

$$\begin{aligned}&= \frac{1}{2}(a - b_{24} + d_4 + d_{4r} - b_{34} - b_{41} - b_{23} + c_v + d_{2r} - e_{1r} - e_2 + d_{3r} + d_2 - e_1 - b_{12} + c_h + d_1 - e_{3r} - e_{2r} + d_3 + d_{1r} - e_3 + c_c - b_{13})\\&= \frac{1}{2}((a) + (-b_{12} - b_{23} - b_{34} - b_{41} - b_{13} - b_{24}) + (c_h + c_v + c_c) + (d_1 + d_2 + d_3 + d_4 + d_{1r} + d_{2r} + d_{3r} + d_{4r}))\\&+ \frac{1}{2}(-e_1 - e_2 - e_3 - e_{1r} - e_{2r} - e_{3r})\end{aligned}$$

as expected.

21+34:

$$-\frac{1}{2}((a)-b_{41}-b_{12}-b_{13}+d_1+d_{1r})$$

12+43: Reflect across vertical axis of 21+43.

$$-\frac{1}{2}(a-b_{23}-b_{12}-b_{24}+d_2+d_{2r})$$

31+24: 21+34 with the two final state particles exchanged

$$-\frac{1}{2}((a)-b_{41}-b_{24}-b_{34}+d_{4r}+d_4)$$

13+42: 31+24 reflected across vertical axis.

$$-\frac{1}{2}(a-b_{23}-b_{13}-b_{34}+d_3+d_{3r})$$

22+33:

$$\frac{1}{2}(a-b_{34}-b_{12}+c_h)$$

23+32: Initial state particles exchanged of 22+33.

$$\frac{1}{2}(a-b_{24}-b_{13}+c_c)$$

14+41:

$$\frac{1}{2}(a-b_{41}-b_{23}+c_v)$$

Add up everything:

Adding up everything, we find all of a, b, and d cancel out. Leftover, we find

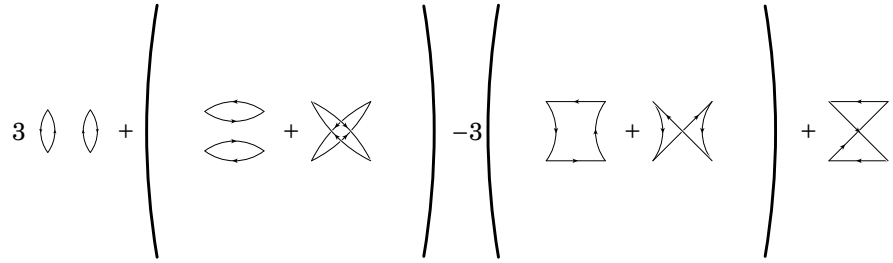
$$(c_v+c_h+c_c)-\frac{1}{2}(e_1+e_2+e_3+e_{1r}+e_{2r}+e_{3r}) \tag{4.14}$$

Total, $I=0$:

Now, we identify the $\frac{1}{3}$ we dropped in section 4.3.3.1 with the 3 on the LHS of eq. (4.15).

$$
\begin{aligned}
3\langle 0|0\rangle &= (c_v + c_h + c_c) - \frac{1}{2}(e_1 + e_2 + e_3 + e_{1r} + e_{2r} + e_{3r}) + (4c_v + 2c_c + 2c_h) \\
&\quad + (-2e_1 - 2e_2 - 2e_{1r} - 2e_{2r}) + 2((2c_v) + (-e_1 - e_2 - e_{1r} - e_{2r} + e_3 + e_{3r})) \\
&= (9c_v + 3c_c + 3c_h) - \frac{9}{2}(e_1 + e_2 + e_{1r} + e_{2r}) + \frac{3}{2}(e_3 + e_{3r}) \\
\Rightarrow \langle 0|0\rangle &= (3c_v + c_c + c_h) - \frac{3}{2}(e_1 + e_2 + e_{1r} + e_{2r}) + \frac{1}{2}(e_3 + e_{3r}) \\
\Rightarrow \langle 0|0\rangle &= (3c_v + c_c + c_h) - 3(e_1 + e_2) + e_3
\end{aligned}
\tag{4.15}
$$

Or, as at the beginning of the section (section 4.3.3):

$\langle I = 0|I = 0\rangle =$ 

4.4 $I = 2$ Contractions

Since Isospin is a good symmetry, we should get the same thing for any bracket $\langle 2, I_3|2, I_3\rangle$ where I_3 is the third component of isospin and $I = 2$ is the total isospin. Thus, we start calculating with $I_3 = 2$ and check with other $I_3 = 1, 0$

4.4.1 $\langle 2,2|2,2\rangle = \langle 2,-2|2,-2\rangle$

$$
\begin{aligned}
&= \langle |\overline{d}\gamma_5 u \overline{d}\gamma_5 u \overline{u}\gamma_5 d \overline{u}\gamma_5 d| \rangle \\
&= (c_h + c_c) - (e_3 + e_{3r})
\end{aligned}
$$

4.4.2 $\langle 2,1|2,1\rangle = \langle 2,-1|2,-1\rangle$

$$
\begin{aligned}
\langle 2,1|2,1\rangle &= \frac{1}{2}(\langle 1,1;1,0|1,1;1,0\rangle + \langle 1,1;1,0|1,0;1,1\rangle + \langle 1,0;1,1|1,0;1,1\rangle + \langle 1,0;1,1|1,1;1,0\rangle) \\
&= \frac{1}{2}\left(\left(\pi^+\pi^0 \to \pi^+\pi^0\right) + \left(\pi^+\pi^0 \to \pi^0\pi^+\right) + \left(\pi^0\pi^+ \to \pi^0\pi^+\right) + \left(\pi^0\pi^+ \to \pi^+\pi^0\right)\right)
\end{aligned}
$$

which is $\frac{1}{2} *$ (4.12 with the first initial and first final particles exchanged+4.12 with the second initial and first final particles exchanged+previous two terms flipped across the horizontal axis.)

This flip across the horizontal axis cancels the $\frac{1}{2}$, and we are left with:

$$\langle 2,1|2,1\rangle = (c_h + c_c) - (e_3 + e_{3r})$$

as expected.

4.4.3 $\langle 2,0|2,0\rangle$

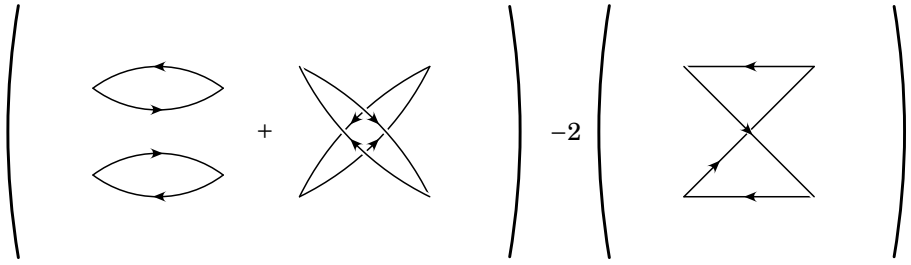
$= \frac{1}{6}(-4*4.12+4.11+4*4.14)$

$$= \frac{1}{6}[-4([2c_v]+[-e_1-e_2-e_{1r}-e_{2r}+e_3+e_{3r}])+(4c_v+2c_c+2c_h)+(-2e_1-2e_2-2e_{1r}-2e_{2r})-2(e_1+e_2+e_3+e_{1r}+e_{2r}+e_{3r})]$$
$$+\frac{1}{6}[4(c_v+c_h+c_c)]$$
$$= \frac{1}{6}[([-8c_v]+4[e_1+e_2+e_{1r}+e_{2r}]-4[e_3+e_{3r}])+(4c_v+2c_c+2c_h)-2(e_1+e_2+e_{1r}+e_{2r})-2(e_1+e_2+e_{1r}+e_{2r})-2(e_3+e_{3r})]$$
$$+\frac{1}{6}[4(c_v+c_h+c_c)]$$
$$= (c_h+c_c)-(e_3+e_{3r})$$

as expected.

4.4.4 Result

Summarizing our findings:

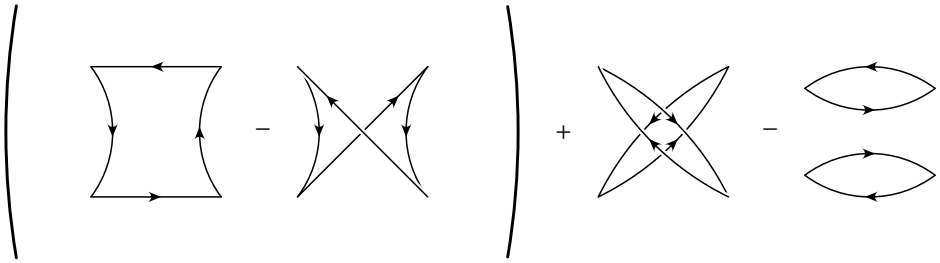
$\langle I=2|I=2\rangle =$ 

4.5 I=1

Similar reasoning gives us $I = 1$. Note that this result comes from $I_3 = 1$, or

$$\pi^0 + \pi^+ \to \pi^0 + \pi^+$$

but, as per usual, is valid for the other I_3's as a result of Isospin symmetry.

$\langle I = 1 | I = 1 \rangle =$ 2 

4.6 Norms of Files vs. Isospin Projections

Here we define the relationship between the contractions in our data sets and the isospin coeffients of chapter 4. As a check, one can compare this document's statements to the isospin coefficients combined with the norm factors in [15] in the file sum_blks.py.

4.6.1 C, D

First, examine section 4.4.4. The first term in parentheses is $2D$ where D refers to our (non-vec) $FigureD$ diagrams (see chapter C). The second term in parentheses is actually equal to the datasets labelled $FigureC$. Why did we choose this normalization? In the production code we averaged over topologies (distinct ways of connecting vertices; e.g. the two diagrams of the first term). We also chose to display the diagrams in chapter 4 such that, in the final results, reversing the particle flow arrows did not result in a new diagram. However, this reversal of arrows constitutes a new topology.

As a check, if we examine section 4.5's last two terms, we find again that these are $2D$ where D is now the vec version. Similarly, if we examine section 4.3.3, the last term is equivalent to $FigureC$.

4.6.2 $R, Vdis, \langle\sigma\rangle$

Next, examine section 4.3.3. The term in the second set of parentheses is $2R$ where R is $FigureR$ (non-vec) as in chapter C. $Vdis$ has no norm factor, so squaring it yields the first term in section 4.3.3 up to the coefficient of 3. Similarly, the scalar bubble $\overline{\psi}\psi$ self-contracted has a norm of 1.

If we examine section 4.5, we see that the first term in parentheses is $-2R$ (vec version) due to the fact that topology differences come with the wrong sign.

4.6.3 T

Figure T has two topologies depending on whether the particle flow is clockwise or counterclockwise. This gives a normalization factor of $\frac{1}{2}$. If we examine section 4.3.2, we see that the triangle shaped diagram inside the parentheses is equivalent to $FigureT$ (scalar), so it needs no norm correction factor.

If we instead examine section 5.3, we see that these are the difference of the two topologies and the topologies are clockwise - counterclockwise (see section 5.3.2), so they similarly need a factor of 2. Thus, e.g., $\langle \pi\pi_{I=1}|\rho\rangle = 2T$ (where T is the vector version).

4.6.4 $pioncorr, pioncorrChk$

The difference in norm is covered in section C.3.2.

4.6.5 $Hbub$

This has norm 1. In fact, the code to compute $pioncorr$ (not $pioncorrChk$) uses the same contraction code as the $Hbub$ diagrams.

Chapter 5

Vector Meson Contractions (ρ)

Starting from Clebsch-Gordan coefficients and the assumption of degenerate mass quarks, we derive contractions for the physical processes $\rho \to \pi\pi$ and $\pi\pi \to \rho$. With the goal of providing suitable projection coefficients for lattice diagrams, we also relate the two sets of contractions under the framework of auxiliary symmetry. These coefficients allow us to start from diagrams where the ρ is only at the source (or sink) and get the diagram with the ρ at the sink (source).

5.1 Introduction

Since we are on the lattice, when we refer to the ρ particle, we do so only implicitly. The actual contractions are for a vector meson with isospin $I = 1$.

We proceed as in the fashion of chapter 4.

5.2 Clebsch-Gordan Coefficients

(For reference on notation, see the beginning of chapter 4). Keeping in mind isopsin rotational symmetry, we calculate contractions for $\rho^+ \to \pi^+\pi^0$ (and the phase convention from section 4.1):

$$
\begin{aligned}
|1,1\rangle &= \frac{1}{\sqrt{2}}(|1,1\rangle|1,0\rangle - |1,0\rangle|1,1\rangle) \\
&= \frac{1}{\sqrt{2}}\left(\pi^+\pi^0 - \pi^0\pi^+\right) \\
&= \frac{1}{2}(\uparrow\uparrow(\uparrow\downarrow+\downarrow\uparrow) - (\uparrow\downarrow+\downarrow\uparrow)\uparrow\uparrow) \\
&= \frac{1}{2}\left(u\overline{d}\left(u\overline{u}-d\overline{d}\right) - \left(u\overline{u}-d\overline{d}\right)u\overline{d}\right)
\end{aligned}
$$

5.3 Contractions

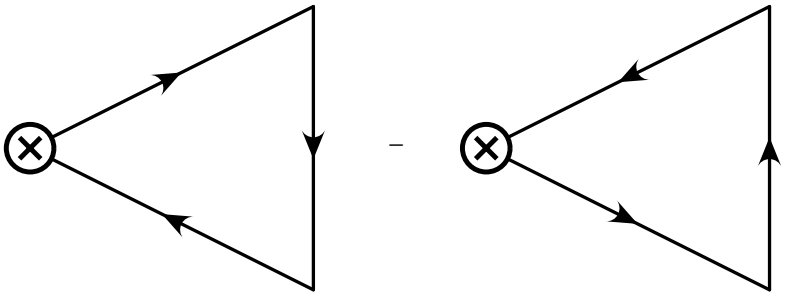
5.3.1 $\pi\pi \to \rho$

Now, because the ρ can't self contract to form a disconnected bubble, we have only two fully connected diagrams. Our correlation functions look like

$$
\begin{aligned}
\langle\rho|\pi\pi\rangle &= \frac{1}{2}\left\langle 0|\overline{d}\gamma^\mu u\left[\overline{d}\gamma^5 u\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right) - \left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\overline{d}\gamma^5 u\right]^\dagger|0\right\rangle \\
&= \frac{1}{2}\left\langle 0|\overline{d}\gamma^\mu u\left[\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\overline{u}\gamma^5 d - \overline{u}\gamma^5 d\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\right]|0\right\rangle
\end{aligned} \tag{5.1}
$$

Clearly we get a triangle shaped diagram subtracted from another triangle diagram with flow going in the opposite direction (the $\frac{1}{2}$ is cancelled by the usual u,d symmetry). In our implementation we refer to these as (vector) T diagrams.

To work out visually how this should appear, we need to establish a convention mapping vertices to fields in our correlation function eq. (5.1). First, we establish time flowing from right to left. Then, we choose the furthest meson to the right as the upper particle (while drawing arrows going from $q \to \overline{q}$). We have

$\langle\rho|\pi\pi_{I=1}\rangle =$  $-$

Importantly, this quantity 2 times what we calculate on the lattice as our figure T (vector).

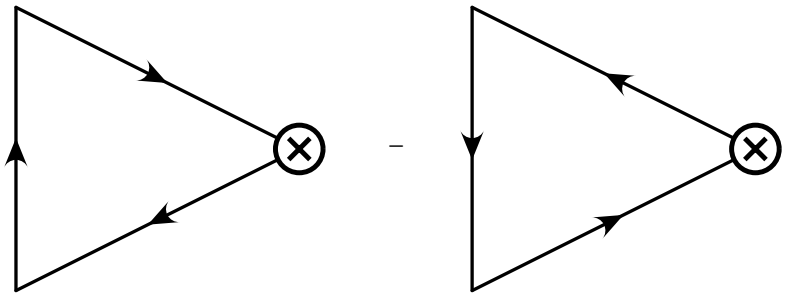
5.3.2 $\rho \to \pi\pi$

$$\begin{aligned}\langle \pi\pi|\rho\rangle &= \frac{1}{2}\left\langle 0\left|\left[\overline{d}\gamma^5 u\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right) - \left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\overline{d}\gamma^5 u\right]\left(\overline{d}\gamma^\mu u\right)^\dagger\right|0\right\rangle \\ &= \frac{1}{2}\left\langle 0\left|\left[\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\overline{d}\gamma^5 u - \overline{d}\gamma^5 u\left(\overline{u}\gamma^5 u - \overline{d}\gamma^5 d\right)\right]\overline{u}\gamma^\mu d\right|0\right\rangle \end{aligned} \tag{5.2}$$

Note in eq. (5.2), we have a negative sign from the dagger of the vector particle under the chiral basis.

We must also establish a visual convention for the particles at the sink (keeping the meaning of the arrows the same as before). We choose the leftmost particle to be the bottom. Thus, we have

$\langle \pi\pi_{I=1}|\rho\rangle =$



Thus, we have the handy mnemonic for remembering these contractions: "clockwise minus counterclockwise."

5.3.3 $\rho \to \rho$

As a check, we should have an overall +1 phase on the $\rho \to \rho$ correlation function (Grassmann cancels the −1 from the † on the source).

5.4 Mapping to A2A Meson Fields

5.4.1 Lattice Vertex Conventions

At this point, we notice that while we have unambiguously defined and calculated our correlation functions, we still have freedom in how these are mapped to meson fields on the lattice. We therefore establish a convention to associate every vertex with a particular time slice (since in general each vertex does occupy a different time slice).

The bottom sink and upper source we define as the inner pions, while the upper sink and bottom source are the outer pions. Visually this means the vertices are

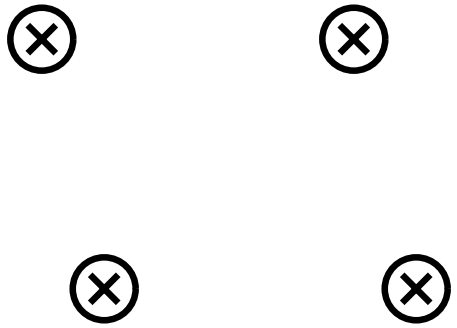


Figure 5.1: Time convention is right to left.

Thus, different diagrams and topologies within these diagrams (or diagram classes) are properly seen as different ways of connecting these temporally fixed vertices.

5.4.2 Auxiliary T diagrams

As in all aux transforms, we

1. swap source and sink quantum numbers
2. swap inner and outer
3. do the proper time slice transform (see the two vertex case in section 2.1.5.3)

Thus, visually, a vertex which starts upper stays upper and vice-versa. This fixes the direction of the pion to pion line prior to doing the aux transform. Thus, a diagram which started clockwise ends up counter clockwise (and vice-versa). We thus conclude that the aux diagram gets an overall minus sign relative to the calculated diagram.

Thus, drawing the mirror diagram is the easiest way to remember the sign of the aux. T.

Chapter 6

Vacuum Subtraction

We present various options (and non-options) for approximating disconnected diagrams in the subtraction of vacuum overlap from connected pieces of disconnected diagrams (called bubbles). Numerical evidence (table 6.1 and chapter F) on a 16^3 lattice (see section 3.1) with heavier than physical pions is presented to justify these choices.

6.1 Introduction

In the study of disconnected diagrams on the lattice, or "vacuum diagrams", we find that the diagram can be split into two separate, fully contracted "bubbles." These bubbles are calculated separately, and then composed into the full diagram time slice by time slice.

Additionally, if one is interested chiefly in vacuum expectation values of operators with non-zero energies, one must then subtract from these diagrams the overlap with the zero energy state ($|0\rangle$), also known as the vacuum. We thus begin our analysis by considering the vacuum expectation value (or "correlation function") defined by two bubble operators, O_1, O_2. We can trivially restate this subtraction in Dirac notation:

$$\begin{aligned}\left\langle 0|O_1O_2^\dagger|0\right\rangle - \langle 0|O_1|0\rangle \left\langle 0|O_2^\dagger|0\right\rangle &= \sum_n \left\langle 0|O_1|n\rangle\langle n|O_2^\dagger|0\right\rangle - \langle 0|O_1|0\rangle \left\langle 0|O_2^\dagger|0\right\rangle \\ &= \sum_{n\neq 0} \left\langle 0|O_1|n\rangle\langle n|O_2^\dagger|0\right\rangle \\ &\equiv \left\langle 0|O_1O_2^\dagger|0\right\rangle - \langle O_1\rangle \left\langle O_2^\dagger\right\rangle\end{aligned}$$

6.2 Theoretical Improvements

For the remainder of the document, we consider bubbles of identical species, so $O_1 = O_2 \equiv O$

6.2.1 Round-off Errors

As a first improvement, we can think about how to do this subtraction on a computer to avoid large roundoff errors. We have the identity

$$\left\langle OO^\dagger \right\rangle - \langle O \rangle \left\langle O^\dagger \right\rangle = \left\langle (O - \langle O \rangle)\left(O^\dagger - \left\langle O^\dagger \right\rangle\right) \right\rangle \tag{6.1}$$

Clearly, for non-negligibly noisy bubbles (as is usually the case), we should expect that the RHS of eq. (6.1) is less prone to roundoff errors that happen when the relative magnitudes of subtracted numbers differ.

6.2.2 Taking the Real Part

If a bubble has no momentum, we should expect that, assuming that taking the complex conjugate of the bilinears which make up O gives no change, the bubble should be real in the continuum. Non-real values can be attributed to lattice artifacts. Hence, we might expect that taking the real part of these zero-momentum bubbles would reduce this noise. This step is usually applied to fully composed diagrams when the imaginary part is dropped, but this step can clearly be applied to bubbles for some possible improvement.

6.2.3 Jackknifing

To generate a jackknife sample, one must leave out a configuration in the ensemble average $\langle ... \rangle$. Thus, because we have two ensemble averages in the RHS of eq. (6.1), we should start at the level of bubbles when leaving out a configuration. So, denoting the ensemble average with the ith configuration left out by $\langle ... \rangle_i$, we should properly have the ith jackknife sample by

$$\left\langle (O - \langle O \rangle_i)\left(O^\dagger - \left\langle O^\dagger \right\rangle_i\right) \right\rangle_i$$

This improvement is unique for this document, in that its justification is not from numerical evidence, but theoretical.

6.2.4 Time Averaging

The bubble is given by a vector of complex numbers of length equal to the time extent of the lattice. Thus, when taking the average $\langle O \rangle$, we get the average for each time slice separately ($\langle O(t) \rangle$). As in section 6.2.2, we should expect in the continuum limit equal averages on each time slice due to the time translational invariance of QCD. Thus, we can take the time average of $\langle O(t) \rangle$ (denoted by $\left\langle \overline{O} \right\rangle$). This fact is denoted $\langle O(t) \rangle = \left\langle \overline{O} \right\rangle$.

While this operation is valid, we should expect that time averaging should not be done because some of the noise is usually correlated. One way to see this is through the following intuition: If we denote ensemble level noise $\epsilon(t)$ (which is ensemble averaged but not time averaged) that goes to zero in the infinite statistics limit, we see that

$$\begin{aligned} O - \langle O(t) \rangle &= O' + \epsilon(t) - \left(\left\langle \overline{O} \right\rangle + \epsilon(t) \right) \\ &= O' - \left\langle \overline{O} \right\rangle \end{aligned}$$

gives better noise than

$$\begin{aligned} O - \left\langle \overline{O} \right\rangle &= O' + \epsilon(t) - \left(\left\langle \overline{O} \right\rangle + \bar{\epsilon} \right) \\ &= O' - \left\langle \overline{O} \right\rangle + (\epsilon(t) - \bar{\epsilon}) \end{aligned}$$

6.2.5 Don't Subtract (Moving Bubbles)

We know that $\langle O \rangle = 0$ if momentum of the bubble is non-zero due to translational symmetry of the action. If we have a situation where $\langle O \rangle$ differs substantially from zero, then we should attempt to do the subtraction to remove the overlap with the vacuum (which gets rid of the $E = 0$ state).

In fact, this option is a specific instance of the general prescription to not do the subtraction at all. Under this regime, $\langle O \rangle \langle O^\dagger \rangle \approx \left\langle \overline{O} \right\rangle \left\langle \overline{O^\dagger} \right\rangle = const.$ We thus can absorb this constant into our overall fit constant (which in the two meson two point case also receives contribution from the so-called around the world term: e.g. one pion travels forwards in time and the other backwards wrapping around the lattice and meeting on the same time slice). Unfortunately, this prescription fails dramatically in generalized eigenvalue analysis. The reason is that a vacuum operator is not included in the matrix of correlation functions, but the corresponding eigenvalue is, leading to (even in the best case), an under-resolved spectrum.

Table 6.1: Comparison of Vacuum Subtraction Options

Option	$t=0$	$t=2$	$t=4$	$t=6$
No sub. ($p=0$) avg$*10^{-10}$, err$*10^{-7}$	1.610621, 8.519775	1.599222, 8.430581	1.595244, 8.379305	1.595097, 8.363323
Take Real (time avg) avg$*10^{-7}$, err$*10^{-6}$	13.77780, 6.217375	2.317461, 5.566338	−1.650743, 5.241048	−1.822182, 5.271675
Take Real (no time avg) avg$*10^{-7}$, err$*10^{-6}$	13.66820, 6.173677	2.312269, 5.547754	−1.616409, 5.251136	−1.713223, 5.254095
Don't Take Real (time avg) real(avg)$*10^{-7}$, err$*10^{-6}$	13.71136, 6.221494	2.312463, 5.561933	−1.664871, 5.237068	−1.812468, 5.273362
Don't Take Real (no time avg) real(avg)$*10^{-7}$, err$*10^{-6}$	13.60080, 6.177820	2.307887, 5.543593	−1.630726, 5.247349	−1.702792, 5.256626
No sub ($p\neq 0$) real(avg)$*10^{-6}$, err$*10^{-6}$	64.99552, 1.290276	23.78021, 1.029827	6.593756, 0.9983585	1.164388, 1.041087
Do Sub. ($p\neq 0$, time avg) real(avg)$*10^{-6}$, err$*10^{-6}$	64.57134, 1.280824	23.35602, 1.019213	6.169573, 0.9931069	0.7402048, 1.041178
Do Sub. ($p\neq 0$, no time avg) real(avg)$*10^{-6}$, err$*10^{-6}$	63.77970, 1.263311	22.95654, 1.009058	6.127332, 0.9923628	0.8991950, 1.039181

We can see small differences, but it is arguable that weakly preferred optimum is to not time average, do subtraction in all cases (including $p\neq 0$), and don't take the real.

6.3 Numerical Results

While these are not totally conclusive (due to constraints of the ensemble and the fact that only one species of bubble is tested), they are very suggestive. We use the ensemble given in section 3.1, with the unthermalized trajectories (config num < 1000) excluded. Full results are given in chapter F.

The noise is measured by jackknife error. The per time slice errors are presented below for the various cases outlined above.

6.4 Conclusion

Final prescription:

1. Use the RHS of eq. (6.1).
2. Don't time average (weakly favored by numerical evidence, examine first/second decimal place in jackknife error).
3. Jackknife at the level of bubbles.

Taking the real part or not taking the real part of zero momentum bubbles seems to have no large effect. Similarly, we get a null result for subtraction/no subtraction of the moving bubble.

Future work might fill in numerical evidence for section 6.2.1, as well as add the disconnected $\langle \pi\pi|\sigma\rangle$ diagrams. We are also cautiously optimistic that new Monte Carlo techniques might be used to rescue the very noisy signals from these disconnected diagrams.

Chapter 7

Angular Momentum

We would like to form operators of definite angular momentum. However, we know the full symmetry group is broken into irreducible representations (irreps) of the discrete rotation/reflection octahedral group. If we are in a moving frame, this symmetry is reduced to helicity states in the corresponding little group[10]. Thus, each irrep of the resulting symmetry group overlaps with a tower of continuum angular momentum states. This correspondence can be found in table II of [1]. Exotic spins (meaning angular momenta higher than the lowest allowed) of the $\pi\pi$ system are beyond the scope of this work; we only look at the lowest spin of each given isospin (consistent with the overall exchange symmetry of a system of two identical bosons) because, anticipating $K \to \pi\pi$, the kaon has $J = 0$.

7.1 Irrep Projections

We focus on the lowest spin for each allowed isospin. In general, our $\pi\pi$ creation operators are defined as (via[1])

$$\pi\pi^{[\vec{k}_1,\vec{k}_2]}_{\vec{P},\Lambda,\mu} = \sum_{\vec{k}_1+\vec{k}_2=\vec{P}} C(\vec{P},\Lambda,\mu;\vec{k}_1,\vec{k}_2)\pi^\dagger(\vec{k}_1)\pi^\dagger(\vec{k}_2)$$

where C is the Clebsch-Gordan coefficient (CG coefficient). The specific CG coefficients for the particular irreps we project onto can be found in the supplementary material of [1]. We know the overall wave function must be symmetric due at least to the fact that they decay from a meson (the kaon). Mesons are composed of two spin $\frac{1}{2}$ particles, and group theory tells us these must add to be an integer spin. Mesons must therefore be bosons from the well-known spin statistics theorem. We thus know the $\pi\pi$ wave function must be overall symmetric. If we examine the lowest compatible spin, we find $I = 0,2$ give $J = 0$ while $I = 1$ gives $J = 1$. We examine

moving frames only for $J = 0$ due to the fact that the section 1.3.1.1 functions are much simpler to code up (in fact we can use the same code for $J = 0, 1$ as long as we are in the center of mass frame). Moving frames for $J = 1$ will be considered in future work. We refer to different rows of an irrep by λ. The dimensionality d of the irrep corresponds to the number of rows. Each row in principle corresponds to a different third component of angular momentum. Thus, in the continuum, the T_1^- rows correspond to different polarizations or the ρ (different directions for its angular momentum vector). The row (singular) of A_1^+ is trivial. Each of these sets of parameters $(I, \vec{P}, \lambda)$ gives a separate GEVP (see chapter 8) for us to solve. It turns out that the one-dimensional irrep A_1^+ corresponds to $J = 0$ (again, see table II **??**), while $d = 3, T_1^-$ to $J = 1$.

Thus, we can specify a GEVP given two parameters:

1. the irrep or the isospin

2. the row of the irrep (if $I = 1$) or $\vec{P}$, the center of mass momentum

7.2 Operator Specification

We can specify an operator in this GEVP (see chapter 8) via the relative momentum between the two pions or the Γ structure (see eq. (2.2)) (e.g., ρ, σ). It turns out (with the exception of $I = 1$, which has a relative minus sign to account for the antisymmetry of $I = 1$) every separate momentum combination (momentum assignment to the individual mesons) is summed with equal weight to make this operator (the absolute value of the $I = 1$ coefficients within an operator is also the same). However, we place a restriction on the inner pions for the moving frames: if one pion in the operator has a greater momentum magnitude, we force it to be the outer pion[1]. One can show this equal weight condition fully specifies the irrep projection.

The lattice is isotropic in all spatial directions (after gauge average), so a given momentum magnitude must give the same energies as another rotated operator (e.g., $\vec{P} = (0, 0, 1)$ gives the same energies as $\vec{P} = (0, 1, 0)$). We thus average these different GEVP matrices. Our max individual particle momenta is $p = (\pm 1, \pm 1, \pm 1)$. We thus calculate up to $(\pm 1, \pm 1, \pm 1)$ for π, ρ, σ.

$I = 1$ has a more complicated phase shift formula for moving frames, so we only look at the stationary frame.

We thus end up with the following 9 separate GEVP's:

[1] It was supposed initially that this condition would, since the inner pion is more energetic and propagates a shorter distance, we would get less noise. However, the fewer momentum combinations entering into these GEVP's seem to reduce statistics in a way which overwhelms the gains of this effect.

Table 7.1: List of separate GEVP's

$\pi\pi$ isospin	$\lvert\vec{P}\rvert$
0	0
0	1
0	2
0	3
1	0
2	0
2	1
2	2
2	3

If a given GEVP has N operators we get N energies and phase shifts (assuming all the states are statistically well-resolved).

7.2.1 Example Operators

Suppose $\vec{P} = 0$ and $I = 0, 2$. For these frames, we have 4 operators (since $|p|_{\max} = 3$):

$$\pi\pi^{|\vec{k}_1|=0}_{\vec{P}=0,A_1^+,0} = \pi^\dagger(0)\pi^\dagger(0)$$

$$\pi\pi^{|\vec{k}_1|=1}_{\vec{P}=0,A_1^+,0} = \sum_{\vec{k}_1} \pi^\dagger(\vec{k}_1)\pi^\dagger(-\vec{k}_1)$$

$$\pi\pi^{|\vec{k}_1|=2}_{\vec{P}=0,A_1^+,0} = \sum_{\vec{k}_1} \pi^\dagger(\vec{k}_1)\pi^\dagger(-\vec{k}_1)$$

$$\pi\pi^{|\vec{k}_1|=3}_{\vec{P}=0,A_1^+,0} = \sum_{\vec{k}_1} \pi^\dagger(\vec{k}_1)\pi^\dagger(-\vec{k}_1)$$

The normalization is not important since the GEVP (see chapter 8) is invariant under an overall rescaling of operator norm.

A full list of the operators we use (including particles and momentum assignments) can be found under utilities/oplist.py in [15].

Chapter 8

Generalized Eigenvalue Problem (GEVP)

Following the lead of [35],[36], we define a matrix of correlation functions $C(t)$ from N operators $\{O_i\}$

$$\begin{aligned} C_{ij}(t = t_{\text{snk}} - t_{\text{src}}) &= \left\langle 0|O_i(t_{\text{snk}})O_j^\dagger(t_{\text{src}})|0\right\rangle \\ &= \sum_{n=0}^{\infty}\left\langle 0|O_i(t_{\text{snk}})|n\right\rangle\left\langle n|O_j^\dagger(t_{\text{src}})|0\right\rangle + \left\langle n|O_i^\dagger(t_{\text{snk}})|0\right\rangle\left\langle 0|O_j(t_{\text{src}})|n\right\rangle \quad (8.1) \\ &= \sum_{n=0}^{\infty}\left\langle 0|O_i(t_{\text{snk}})|n\right\rangle\left\langle n|O_j^\dagger(t_{\text{src}})|0\right\rangle + (c.c., t \to -t) \quad (8.2) \\ &= \sum_{n=0}^{\infty}\left(\psi_{in}\psi_{jn}^* e^{-E_n t} + \psi_{in}^*\psi_{jn} e^{-E_n(L_t - t)}\right) \quad (8.3) \\ \psi_{in} &\equiv \langle 0|O_i|n\rangle \end{aligned}$$

Clearly, the $C(t)$ is hermitian[1]. Assume $\left\langle\left(\psi_{in}\psi_{jn}^*\right)^*\right\rangle = \left\langle\psi_{in}\psi_{jn}^*\right\rangle$ since this is demonstrably true of all our operators except for eqs. (5.1) and (5.2) which are pure imaginary and can be made pure real by redefining $\overline{\psi}\gamma_\mu\psi \to i\overline{\psi}\gamma_\mu\psi$ which we do in our analysis code. Furthermore, we know the second term in eqs. (8.1) and (8.2) can be justified (up to an overall normalization factor) via crossing symmetry. These terms come from propagation through the boundary and include the other exponential in the cosh in eq. (8.3) as well as around the world terms we've neglected for now (see section 8.4). Our correlation function data is complex, but we know

[1] We do in fact calculate both C_{ij} and C_{ji} (to achieve perhaps some safety via redundancy and hypothetically better statistics, although this latter property hasn't been checked). We then enforce hermiticity by transforming $C(t) \to \frac{1}{2}\left(C(t) + C^\dagger(t)\right)$.

the imaginary parts are due to A2A noise (see section 2.0.1). Then,

$$C_{ij}(t) = \sum_{n=0}^{\infty} \psi_{in}\psi_{jn}^{*}\left(e^{-E_n t} + e^{-E_n(L_t - t)}\right)$$

Now, we define a generalized eigenvalue problem (GEVP):

$$C(t)v_n(t,t_0) = \lambda_n(t,t_0)C(t_0)v_n(t,t_0) \tag{8.4}$$

If we find u_n such that $\langle \psi_m | u_n \rangle = \sum_{i=0}^{N-1} u_{ni}\psi_{mi}^{*} = \delta_{nm}$ then

$$\begin{aligned} C(t)v_n(t,t_0) &= \sum_{m=0}^{\infty} \psi_{im}\psi_{jm}^{*}u_{jn}\left(e^{-E_m t} + e^{-E_m(L_t - t)}\right) \\ &= \sum_{n=0}^{\infty} \psi_{im}\delta_{nm}\left(e^{-E_m t} + e^{-E_m(L_t - t)}\right) \\ &= \psi_{in}\left(e^{-E_n t} + e^{-E_n(L_t - t)}\right) \end{aligned}$$

So $\lambda_n = \frac{e^{-E_n t_0} + e^{-E_n(L_t - t_0)}}{e^{-E_n t} + e^{-E_n(L_t - t)}}, v_n = u_n$ solves the GEVP (eq. (8.4)). However, because we N operators, we can only get at most N solutions to eq. (8.4). Suppose only N states overlap our operators, though. In this case, solving the GEVP gets us eigenstates of our QCD Hamiltonian and eigenvalues which can be manipulated to find energies. Furthermore, because noise in lattice QCD is multiplicative, solving the GEVP usually reduces the noise that we might find if we only diagonalized one of the GEVP matrices ($C(t), C(t_0)$. Finally, the GEVP framework allows us to make precise quantitative predictions as to the asymptotic bounds on the systematic error from excited states.

For these reasons, we choose to solve a GEVP before we do any fits.

In the remaining sections of this chapter, we collect some theorems that allow us to refine our GEVP spectra.

8.1 First Order Excited State Systematic Error

Let $A \equiv C(t), B \equiv C(t_0)$. Then we can define our GEVP by

$$A|v\rangle = \lambda B|v\rangle$$
$$B^{-1}A|v\rangle = \lambda|v\rangle$$
$$\sim \frac{A}{B}|v\rangle = \lambda|v\rangle$$

At zeroth order we assume A,B with dimension N contain N distinct states. A_1,B_1 contain the rest of the states, so we have, e.g., $A = A_0 + A_1$.

We now prove a weaker form of the well-known theorem from [35] (their argument strengthens the $t - t_0 = const$ condition to $t_0 \geq \frac{t}{2}$).

Theorem 1. *The first order*[2] *correction due to excited states to energy E_m of the GEVP is $O(e^{-(E_{N+1}-E_m)t})$ provided that $t - t_o = const$.*

Proof. We don't care about matrix orders, just the powers of the matrices. We want to find λ_1 (For λ_1,E_1, the first order is defined via size instead of powers of A_1,B_1). Collect what we know about the individual matrix sizes (meaning asymptotic in t,t_0) and the sizes of the perturbations:

$$v_{m,0} = O(1)$$
$$A_0 v_{m,0} \sim e^{-E_m t}$$
$$B_0 v_{m,0} \sim e^{-E_m t_0}$$
$$A_1 \sim e^{-E_{N+1} t}$$
$$B_1 \sim e^{-E_{N+1} t_0}$$
$$\lambda_0 \sim e^{-E_m (t-t_0)}$$
$$\frac{1}{\lambda_0} \sim \frac{B_0}{A_0}$$
$$\frac{1}{\lambda_0}\left(\frac{A}{B}\right)_1 = \frac{A_1}{A_0} - \frac{B_1}{B_0} + O\left(\frac{B_1^2}{B_0^2}\right) \tag{8.5}$$

Now, assume $\log\left(\frac{A}{B}\right)$ is well-defined (this assumption is not strictly necessary, one could also find $E_1 = -\frac{\lambda_0' + \lambda_1'}{\lambda_0}(1 - \frac{\lambda_1}{\lambda_0})$ which one can show does give the same results. The expansion of the log is the same as the expansion of the perturbation series, so the key term eq. (8.8) comes from $\frac{1}{\lambda}$). We can expand the infinite series eq. (8.5) in $\frac{B_1}{B_0}$ or we can expand the log power series in powers of $\frac{1}{\lambda_0} \sim \frac{B_0}{A_0}$. If we do both to sufficient

[2] in size, as opposed to power of A_1,B_1

order, we should be able to truncate and locate the expressions which give first order corrections to the energy.

$$\log(\lambda)_1 \sim \frac{A_1}{A_0} - \frac{B_1}{B_0} + \left(\frac{B_1}{B_0}\right)^2 + O\left(\frac{A_1^2}{A_0^2}\frac{B_0}{A_0}, \frac{A_1 B_1}{A_0 B_0}\frac{B_0}{A_0}, ...\right) \tag{8.6}$$

where ... denotes higher order.

$\left(\frac{B_1}{B_0}\right)^n$ in eq. (8.5) has no t dependence, so these will vanish under $-\partial_t$. N.B. If we don't take the derivative (especially if $t - t_0 \neq const.$) and only divide by $t - t_0$, we will get less suppression of the excited states (t_0 dependence instead of t). Let's examine the two relevant remaining terms in eq. (8.6): The first $\frac{A_1}{A_0}$ is the usual quoted suppression factor.

$$\frac{A_1}{A_0} \sim e^{-(E_{N+1} - E_m)t} \tag{8.7}$$

The second is

$$\left(\frac{B_1}{B_0}\right)^2 \tag{8.8}$$

This term is more complicated, and in fact is the origin of the strengthened t, t_0 condition $t_0 \geq \frac{t}{2}$. (The 2 clearly comes from the 2 in the LHS of eq. (8.8).) For $t - t_0 = const.$, it at least has a term that is $O(e^{-(E_{N+1} - E_n)t_0})$ since (we know from usual perturbation theory) it comes from expanding the eigenvector and this eigenvector expansion multiplies a term which is already at least that small. (As an aside we need to assume $E_{N+1} - E_m = O(1)$ so that ∂_t does not change appreciably the size of the coefficient of the systematic error.)

Thus, the largest term in eq. (8.6) is $O(e^{-(E_{N+1} - E_n)t_0})$ but $O(e^{-(E_{N+1} - E_n)t_0}) = O(e^{-(E_{N+1} - E_n)t})$.

□

8.2 Explicit GEVP Perturbation Theory

We redo the perturbation theory in section 8.1 explicitly (not counting powers) to examine two special cases (sections 8.2.2 and 8.2.3).

For convenience, identify $A \equiv C(t), B \equiv C(t_0)$. Then expand to first order in α to find $\left|v^1\right\rangle$ (following the analogous usual steps from perturbation theory)

$$
\begin{aligned}
Av &= \lambda Bv \\
(A^0 + A^1)(v^0 + v^1) &= (\lambda^0 + \lambda^1)(B^0 + B^1)(v^0 + v^1) + O(\alpha^2) \\
\Rightarrow A^0 v^1 + A^1 v^0 &\approx \lambda^0 B^0 v^1 + \lambda^1 B^0 v^0 + \lambda^0 B^1 v^0 \\
\Rightarrow (A^0 - \lambda^0 B^0) v^1 &\approx \left(-A^1 + \lambda^1 B^0 + \lambda^0 B^1\right) v^0
\end{aligned}
$$

Using the fact that $\{v_m^0\}$ constitute a basis for the GEVP eigenspace, $v_m^0 A^0 v_n^0 \sim \delta_{mn}, v_m^0 B^0 v_n^0 \sim \delta_{mn}$, and using the argument from perturbation theory to eliminate λ^1, we find

$$
\begin{aligned}
\Rightarrow \sum_{m \neq n} (A^0 - \lambda_n^0 B^0) v_m^0 c_{mn} &\approx \left(-A^1 + \lambda^1 B^0 + \lambda_n^0 B^1\right) v_n^0 \\
\Rightarrow c_{mn} &\approx \frac{\lambda_n^0}{\left(\lambda_m^0 - \lambda_n^0\right)} \left(\frac{\langle v_m^0 | B^1 | v_n^0 \rangle}{\langle v_m^0 | B^0 | v_m^0 \rangle} - \frac{\lambda_m^0}{\lambda_n^0} \frac{\langle v_m^0 | A^1 | v_n^0 \rangle}{\langle v_m^0 | A^0 | v_m^0 \rangle} \right)
\end{aligned}
$$

$$
\begin{aligned}
&A |v\rangle = \lambda B |v\rangle \\
&\Rightarrow \lambda = \frac{\langle v | A | v \rangle}{\langle v | B | v \rangle} \\
&\quad = \frac{\langle v^0 | A^0 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \frac{\langle v^0 | A^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} - \frac{\langle v^0 | A^0 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \frac{\langle v^0 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \\
&\quad + \frac{\langle v^0 | A^0 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \left(\frac{\langle v^0 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \right)^2 - \frac{\langle v^0 | A^0 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \frac{\langle v^0 | B^1 | v^1 \rangle}{\langle v^0 | B^0 | v^0 \rangle} - \frac{\langle v^0 | A^0 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \frac{\langle v^1 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \\
&\quad + \frac{\langle v^0 | A^1 | v^1 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \frac{\langle v^1 | A^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + O(\alpha^3) \\
&\quad = \lambda_n^0 \left(1 + \frac{\langle v^0 | A^1 | v^0 \rangle}{\langle v^0 | A^0 | v^0 \rangle} - \frac{\langle v^0 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \left(\frac{\langle v^0 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} \right)^2 - \frac{\langle v^0 | B^1 | v^1 \rangle}{\langle v^0 | B^0 | v^0 \rangle} - \frac{\langle v^1 | B^1 | v^0 \rangle}{\langle v^0 | B^0 | v^0 \rangle} + \frac{\langle v^0 | A^1 | v^1 \rangle}{\langle v^0 | A^0 | v^0 \rangle} + \frac{\langle v^1 | A^1 | v^0 \rangle}{\langle v^0 | A^0 | v^0 \rangle} \right) + O(\alpha^3)
\end{aligned}
$$

8.2.1 Generalized Orthogonality Relation

Assume non-degenerate spectrum ($\lambda_n \neq \lambda_m$). Assume as well that we have $N \times N$ hermitian GEVP matrices $C(t)$ at three separate time slices.

$$
\begin{aligned}
C(t_0)|v_n\rangle &= \lambda_n C(t)|v_n\rangle \\
\langle v_m|C(t_0)^\dagger &= \lambda_m^* \langle v_m|C(t)^\dagger \\
\langle v_m|C(t_0) &= \lambda_m \langle v_m|C(t) \\
\langle v_m|C(t_0)|v_n\rangle &= \lambda_n \langle v_m|C(t)|v_n\rangle \\
&= \lambda_m \langle v_m|C(t)|v_n\rangle \\
\Rightarrow 0 &= (\lambda_n - \lambda_m)\langle v_m|C(t)|v_n\rangle \\
\Rightarrow 0 &= \langle v_m|C(t)|v_n\rangle
\end{aligned} \tag{8.9}
$$

We can define perturbation theory via

$$
\begin{aligned}
C(t) &= C^0(t) + C^1(t) \\
C^0_{ij}(t) &= \sum_{n=1}^{N} \psi_{in}\psi^*_{jn} e^{-E_n t} \\
C^1_{ij}(t) &= \sum_{n=N+1}^{\infty} \psi_{in}\psi^*_{jn} e^{-E_n t}
\end{aligned}
$$

Now, expand eq. (8.9) out to to first order in C. If we use some parameter α to count powers, we can take derivatives of α and evaluate at $\alpha = 0$ which demonstrates that eq. (8.9) is true order by order in C.

$$
\begin{aligned}
\left\langle v_m(t,t_0)|C(t')|v_n(t,t_0)\right\rangle = &\left\langle v^0_m(t,t_0)|C^0(t')|v^0_n(t,t_0)\right\rangle + \\
&+ \left\langle v^1_m(t,t_0)|C^0(t')|v^0_n(t,t_0)\right\rangle + \left\langle v^0_m(t,t_0)|C^0(t')|v^1_n(t,t_0)\right\rangle + \left\langle v^0_m(t,t_0)|C^1(t')|v^0_n(t,t_0)\right\rangle + O(\alpha^2)
\end{aligned} \tag{8.10}
$$

Now, we know that $\left|v^0\right\rangle \equiv \left|v^0(t,t_0)\right\rangle = \left|v^0(t',t_0)\right\rangle$ since the zeroth order eigenvector is time independent.

$$
\begin{aligned}
\Rightarrow \left\langle v^0_m(t,t_0)|C^0(t')|v^0_n(t,t_0)\right\rangle &= \left\langle v^0_m|C^0(t')|v^0_n\right\rangle \\
&= \left\langle v^0_m|C^0(t_0)|v^0_n\right\rangle \lambda(t',t_0) \\
&= 0
\end{aligned} \tag{8.11}
$$

In other words, we show in eq. (8.11) that eq. (8.10) is 0 at zeroth order (again, since we know the orthogonality

relation should hold order by order in our power counting expansion α). Then,

$$\begin{aligned}
\Rightarrow \langle v_m(t,t_0)|C(t')|v_n(t,t_0)\rangle &= \langle v_m^1(t,t_0)|C^0(t')|v_n^0\rangle + \langle v_m^0|C^0(t')|v_n^1(t,t_0)\rangle + \langle v_m^0|C^1(t')|v_n^0\rangle + O(\alpha^2) \\
&\approx \lambda_n^0(t',t)\langle v_m^1(t,t_0)|C^0(t)|v_n^0\rangle + \lambda_m^0(t',t)\langle v_m^0|C^0(t)|v_n^1(t,t_0)\rangle + \langle v_m^0|C^1(t')|v_n^0\rangle \\
&= \lambda_n^0(t',t)\langle v_m^1(t,t_0)|A^0|v_n^0\rangle + \lambda_m^0(t',t)\langle v_m^0|A^0|v_n^1(t,t_0)\rangle + \langle v_m^0|C^1(t')|v_n^0\rangle \\
&= \frac{\langle v_m^0|B^1|v_n^0\rangle \lambda_m^0\lambda_n^0}{(\lambda_n^0-\lambda_m^0)}\left(\lambda_n^0(t',t)-\lambda_m^0(t',t)\right) - \frac{\langle v_m^0|A^1|v_n^0\rangle}{(\lambda_n^0-\lambda_m^0)}\left(\lambda_n^0(t',t_0)-\lambda_m^0(t',t_0)\right) + \langle v_m^0|C^1(t')|v_n^0\rangle
\end{aligned} \tag{8.12}$$

As $t' \to t$ or $t' \to t_0$, the RHS of eq. (8.12) goes to 0, as expected. eq. (8.12) has convergence properties that are not easily determined. For that reason, we should look at the other version of eq. (8.12) where the matrix time slice is shared by one of the eigenvectors (assume $t_0 < t \le t_a, t_b$):

$$\begin{aligned}
\langle v_m(t,t_0)|C(t_0)|v_n(t_a,t_b)\rangle &= \langle v_m^1(t,t_0)|C^0(t_0)|v_n^0\rangle + \langle v_m^0|C^0(t_0)|v_n^1(t_a,t_b)\rangle + \langle v_m^0|C^1(t_0)|v_n^0\rangle + O(\alpha^2) \\
&\approx \langle v_m^1(t,t_0)|C^0(t_0)|v_n^0\rangle + \langle v_m^0|C^0(t_0)|v_n^1(t_a,t_b)\rangle + \langle v_m^0|C^1(t_0)|v_n^0\rangle \\
&= \langle v_m^0|C^0(t_0)|\left(|v_n^1(t_a,t_b)\rangle - |v_n^1(t,t_0)\rangle\right) \\
&\approx -\langle v_m^0|C^0(t_0)|v_n^1(t,t_0)\rangle \\
&= \frac{\langle v_m^0|C^0(t_0)|v_m^0\rangle}{\lambda_n^0-\lambda_m^0}\left(\lambda_n^0\frac{\langle v_m^0|B^1|v_n^0\rangle}{\langle v_m^0|B^0|v_m^0\rangle} - \lambda_m^0\frac{\langle v_m^0|A^1|v_n^0\rangle}{\langle v_m^0|A^0|v_m^0\rangle}\right) \\
\Rightarrow \frac{\langle v_m(t,t_0)|C(t_0)|v_n(t_a,t_b)\rangle}{\langle v_m(t,t_0)|C(t_0)|v_m(t,t_0)\rangle} &\approx \frac{1}{\lambda_n^0-\lambda_m^0}\left(\lambda_n^0\frac{\langle v_m^0|B^1|v_n^0\rangle}{\langle v_m^0|B^0|v_m^0\rangle} - \lambda_m^0\frac{\langle v_m^0|A^1|v_n^0\rangle}{\langle v_m^0|A^0|v_m^0\rangle}\right) \qquad (8.13) \\
&\approx \frac{\lambda_n^0}{\lambda_n^0-\lambda_m^0}\frac{\langle v_m^0|B^1|v_n^0\rangle}{\langle v_m^0|B^0|v_m^0\rangle} \qquad (8.14)
\end{aligned}$$

where in eq. (8.14) we've dropped all but the leading term.

Otherwise, we can use this relation to help eliminate systematic error since we know the LHS and all the terms in eq. (8.12) appear in perturbative expansion of the GEVP. This expansion is also suggestive since, as we note in eq. (8.30), the eigenvectors we use need not actually solve the GEVP for A, B. Indeed, for p time slices, we have $\binom{p}{2}$ eigenvectors for each dimension of our GEVP defined by their time pairing. We also know

$$\langle v_n(t_1,t_2)|C(t)|v_n(t_1,t_2)\rangle + O(\alpha) = \lambda_n \in \mathbb{R} \tag{8.15}$$

We have p choices for the matrix $C(t)$ and $\binom{\binom{p}{2}}{2} + \binom{p}{2} = \frac{1}{8}(p^2-p)^2 + \frac{1}{4}(p^2-p)$ choices for the eigenvectors. We thus have $\frac{p}{8}(p^2-p)^2 + \frac{p}{4}(p^2-p)$ ways to form eq. (8.15). In fact, we also have neglected the freedom to allow the

times to vary of the eigenvectors of eq. (8.12), so we get even more equations to work with. While nominally this gives us many more estimates for λ, most are likely to be noisy even if they give different systematic errors. Since they depend on the same time slice data, it is likely that averaging them is not going to help statistics. eq. (8.14) has the nice feature that it is pure systematic, so all the errors are not overwhelmed by some large λ we need to subtract. Moreover, the division in eq. (8.13) likely will remove some noise due to most of its components being identical to the numerator.

8.2.2 Weakly Coupled Excited States (or diagonal C^1 in the $\{v_n^0\}$ basis)

We suppose

$$\langle v_m^0 | C^1 | v_n^0 \rangle \sim \delta_{mn}$$
$$\Rightarrow v^1 \approx 0$$

and that C_{ii}^1 is described by a single state for all $1 < i \leq N$.

Then,

$$\begin{aligned}
a_n e^{-\Delta_n t} &\equiv \frac{\langle v_n^0 | A_{nn}^1 | v_n^0 \rangle}{\langle v_n^0 | A^0 | v_n^0 \rangle} \\
\Rightarrow a_n e^{-\Delta_n t_0} &\approx \frac{\langle v_n^0 | B^1 | v_n^0 \rangle}{\langle v_n^0 | B^0 | v_n^0 \rangle} \\
\Rightarrow \lambda_n &\approx \lambda_n^0 \left(1 + a_n e^{-\Delta_n t} - a_n e^{-\Delta_n t_0} + a_n^2 e^{-2\Delta_n t_0} - \frac{\langle v_n^0 | B^1 | v_n^1 \rangle}{\langle v_n^0 | B^0 | v_n^0 \rangle} - \frac{\langle v_n^1 | B^1 | v_n^0 \rangle}{\langle v_n^0 | B^0 | v_n^0 \rangle} + \frac{\langle v_n^0 | A^1 | v_n^1 \rangle}{\langle v_n^0 | A^0 | v_n^0 \rangle} + \frac{\langle v_n^1 | A^1 | v_n^0 \rangle}{\langle v_n^0 | A^0 | v_n^0 \rangle} \right) + O(\alpha^3) \\
&\approx \lambda_n^0 \left(1 + a_n e^{-\Delta_n t} - a_n e^{-\Delta_n t_0} + a_n^2 e^{-2\Delta_n t_0} \right) + O(\alpha^3) \\
&= \lambda_n^0 \left(1 + a_n e^{-\Delta_n t} + \frac{a_n e^{-\Delta_n t_0}}{1 + a_n e^{-\Delta_n t_0}} \right) \qquad (8.16)
\end{aligned}$$

Note that $a_n^2 = a_n a_n$ (it is a square, not a power count in α). Now, assuming we have three time slices for $C(t)$, we can form $\binom{3}{2} = 3$ GEVP's which allow us to solve numerically for the three unknowns in eq. (8.16): $\lambda_n^0, \Delta_n, a_n$.

8.2.3 Well Separated Excited States

We suppose

$$A^1_{ij} \approx e^{-E_{N+1}t}\psi_{iN}\psi^*_{jN}$$
$$B^1_{ij} \approx e^{-E_{N+1}t_0}\psi_{iN}\psi^*_{jN}$$

Then,

$$\begin{aligned}
\Delta_n &\equiv E_{N+1} - E_n \\
\lambda_n(t,t_0) &= e^{-E_n(t-t_0)} \\
a_n e^{-\Delta_n t} &\equiv \frac{\langle v_n^0|A^1|v_n^0\rangle}{\langle v_n^0|A^0|v_n^0\rangle} \\
\Rightarrow a_n e^{-\Delta_n t_0} &\approx \frac{\langle v_n^0|B^1|v_n^0\rangle}{\langle v_n^0|B^0|v_n^0\rangle} \\
\Rightarrow \lambda_n^0 &\approx \lambda_n^0\left(1 + a_n e^{-\Delta_n t} - a_n e^{-\Delta_n t_0} + a_n^2 e^{-2\Delta_n t_0} - \frac{\langle v_n^0|B^1|v_n^1\rangle}{\langle v_n^0|B^0|v_n^0\rangle} - \frac{\langle v_n^1|B^1|v_n^0\rangle}{\langle v_n^0|B^0|v_n^0\rangle} + \frac{\langle v_n^0|A^1|v_n^1\rangle}{\langle v_n^0|A^0|v_n^0\rangle} + \frac{\langle v_n^1|A^1|v_n^0\rangle}{\langle v_n^0|A^0|v_n^0\rangle}\right) + O(a^3)
\end{aligned}$$

Now we have the terms in brackets to deal with. Take as an example the first term:

$$\begin{aligned}
\frac{\langle v_n^0|B^1|v_n^1\rangle}{\langle v_n^0|B^0|v_n^0\rangle} &= \sum_{m\neq n} \frac{\langle v_n^0|B^1|v_m^0\rangle}{\langle v_n^0|B^0|v_n^0\rangle} \frac{1}{(\lambda_m^0 - \lambda_n^0)}\left(\lambda_n^0 \frac{\langle v_m^0|B^1|v_n^0\rangle}{\langle v_m^0|B^0|v_m^0\rangle} - \lambda_m^0 \frac{\langle v_m^0|A^1|v_n^0\rangle}{\langle v_m^0|A^0|v_m^0\rangle}\right) \\
b_{mn}e^{-\Delta_n t} &\equiv \frac{\langle v_m^0|A^1|v_n^0\rangle}{\langle v_n^0|A^0|v_n^0\rangle} \\
b_{mn}e^{-\Delta_n t_0} &\approx \frac{\langle v_m^0|B^1|v_n^0\rangle}{\langle v_n^0|B^0|v_n^0\rangle} \qquad (8.17) \\
\Rightarrow \frac{\langle v_n^0|B^1|v_n^1\rangle}{\langle v_n^0|B^0|v_n^0\rangle} &\approx \sum_{m\neq n} b^*_{mn} e^{-\Delta_n t_0} \frac{1}{(\lambda_m^0 - \lambda_n^0)}\left(\lambda_n^0 b_{mn} e^{-\Delta_n t_0} - \lambda_m^0 b_{mn} e^{-\Delta_n t}\right) \\
&= \sum_{m\neq n} \frac{|b_{mn}|^2}{(\lambda_m^0 - \lambda_n^0)}\left(\lambda_n^0 e^{-(\Delta_n+\Delta_m)t_0} - \lambda_m^0 e^{-\Delta_n(t+t_0)}\right)
\end{aligned}$$

In section 8.2.2, we could solve each of the eigenvalues separately. In section 8.2.3, we must solve them simultaneously. We get $3N$ equations, and our unknowns are $b_{mn}, E_n, E_{N+1}, a_n$

How many unknowns are there? We can solve by induction. If $N = 1$ then we get 3 equations with unknowns E_0, a_0, E_{N+1}. Adding a second operator we have another 3 equations and unknowns E_1, a_1, b_{21}. Adding an operator to $N-1$ other operators we get $2+(N-1)$: E_N, a_N, b_{Ni} where $i \neq N$. Summing, we find the number of underdetermined coefficients is $\sum_{n=2}^{N}(n-2) = \frac{1}{2}(N^2 - 3N + 2)$. We thus can exactly solve only the $N = 2$ case if we have 3 time slices. It can be shown that if we have N time slices we can exactly solve an $N-1$

dimensional GEVP.

We could of course also substitute eq. (8.14) in eq. (8.17). Using any relation from section 8.2.1 would of course allow us to completely solve the system (numerically), but the usefulness of these untested techniques likely are very dependent on the numeric/statistical properties of the system being analyzed. However, if eq. (8.14) is resolvable, this eliminates error up to and including size $O(e^{-\Delta t})$, which is an improvement over section 8.1.

8.3 Matrix Elements

We can also use the GEVP to make statements about the excited state error on matrix elements. We find the asymptotic systematic error size of the three-point GEVP. We generally neglect vector structure (including the ordering of the matrices) since the eigenvectors are constant at first order in the expansion of the matrix element.

8.3.1 3pt GEVP, Asymmetric Case

This case requires us to pay closer attention to the vector structure (unlike the symmetric case), but we examine this case because the problem should have a sensible symmetric limit and is a slightly more informative exercise than the symmetric case.

We define the GEVP on a matrix $C(t)$ hermitian at $O(\epsilon)$ using notation of [37]. We assume for all t that $v_n C(t) v_n \neq 0$ (see section 8.6).

$$
\begin{aligned}
C(t,\epsilon) &= \begin{pmatrix} C^{(B)}(t) & \epsilon K(t) \\ \epsilon K^\dagger(t) & D^{(A)}(t) \end{pmatrix} + O(\epsilon^2) \\
C(t) v_{nm} &= \lambda C(t_0) v_{nm} \\
\Rightarrow v_{nm}^\dagger C(t) &= \lambda v_{nm}^\dagger C(t_0) \\
v_{nm} &= \{w_n, u_m\} \\
v_{nm}^\dagger C^{-1}(t_0) C(t) v_{nm} &= \lambda v_{nm}^\dagger v_{nm} \\
\frac{v_{nm}^\dagger C(t) v_{nm}}{v_{nm}^\dagger C(t_0) v_{nm}} &= \lambda
\end{aligned}
\tag{8.18}
$$

Now, we drop the A,B labels since they are redundant with C,D. We also use subscripts to indicate time dependence: (e.g.) $M \equiv M(t)$, and $M_0 \equiv M(t_0)$. Assume we've normalized the eigenvectors: $w_n^\dagger w_n = u_m^\dagger u_m = 1, \Rightarrow v_n^\dagger v_n = 2$. The problem separates into two independent equations just as in A.9 in [37] for the two components

of v_{nm} in eq. (8.18). Thus, we can ignore the $K^\dagger$ terms.

$$\begin{aligned}
\lambda &= \frac{w_n C w_n + u_m D u_m + w_n K u_m + u_m K^\dagger w_n}{w_n C_0 w_n + u_m D_0 u_m + w_n K_0 u_m + u_m K_0^\dagger w_n} \\
&= \frac{w_n C w_n + u_m D u_m}{w_n C_0 w_n + u_m D_0 u_m}\left(1 + \frac{w_n K u_m + u_m K^\dagger w_n}{w_n C w_n + u_m D u_m} - \frac{w_n K_0 u_m + u_m K_0^\dagger w_n}{w_n C_0 w_n + u_m D_0 u_m}\right) + O(\epsilon^2) \\
\frac{d}{d\epsilon}\log(\lambda) &= \frac{w_n K u_m + u_m K^\dagger w_n}{w_n C w_n + u_m D u_m} - \frac{w_n K_0 u_m + u_m K_0^\dagger w_n}{w_n C_0 w_n + u_m D_0 u_m} \\
M_{mn} &= \partial_t\left(\frac{w_n K_0 u_m + u_m K_0^\dagger w_n}{w_n C_0 w_n + u_m D_0 u_m} - \frac{w_n K u_m + u_m K^\dagger w_n}{w_n C w_n + u_m D u_m}\right) \\
M_{mn} &= 2Re\left[\partial_t\left(\frac{w_n K_0 u_m}{w_n C_0 w_n + u_m D_0 u_m} - \frac{w_n K u_m}{w_n C w_n + u_m D u_m}\right)\right] \\
\Gamma(t) &\equiv O(max\{e^{-\Delta_A t}, e^{-\Delta_B t}\})
\end{aligned} \tag{8.19}$$

Now we are concerned with expanding our result eq. (8.19) to first order. We assume that the first order expansion of the three point correlation function matrix is much smaller than the first order expansions in the two point correlation function matrices. We thus, as indicated above, simply insert the expansions from section 8.1. The expansions of sectors A,B are in principle independent at first order. A ratio of $\frac{K}{C,D}, \frac{K_0,K_0^\dagger}{C_0,D_0}$ always gives $\Gamma(t),\Gamma(t_0)$ respectively. It's not too difficult to then see that we end up with first order systematic error which is of size $\partial_t[\Gamma(t)\cdot max\{K,K_0\}]$. At zeroth order in the above expansion we find the summation in the formula for K gives us a factor of t,t_0 for K,K_0 respectively. Thus, the first order systematic error is contained in $\partial_t[t\Gamma(t)]$. After differentiating, we find our final error term

$$\boxed{M^1_{mn} = min(\Delta_A,\Delta_B)t\Gamma(t)} \tag{8.20}$$

So we see the first order systematic error in M_{mn} is the max of that generated by C,D. That means the error is generated by the lowest excited state we leave out of either the (A) or (B) sector. This reduces to

$$M^1_{mn} = \Delta t\Gamma(t)$$

in the symmetric limit, as expected (from the symmetric case in [37]). In [37], they find $\Gamma(t)\to\Gamma(t_0)$ in the asymmetric case (compared to the symmetric). While this expression has a worse convergence rate and no clear way to take the symmetric limit (making its connection to the symmetric case unclear), we should try to understand why it is different. Essentially, if we focus only on matrix powers and replace the sums in the denominators of eq. (8.19) with the bracket which gives larger errors, their expressions are (roughly) related to eq. (8.19) by factors of $\sqrt{\frac{C_0}{D_0}}$ or $\sqrt{\frac{D_0}{C_0}}$. These clearly generate $\Gamma(t_0)$ error terms.

However, even then we get an error term of $\partial_t(t\Gamma(t_0)) = \Gamma(t_0)$. Indeed, it is hard to see how we could find a term $f(t,t_0)$ such that

$$\partial_t f(t,t_0) = \Delta t \Gamma(t_0)$$

since we only have terms which are linear in t, t_0 (due to the 3pt sums) and terms which are exponential in t, t_0. Thus, if we are supposed to derive a term propartional to t (or Δ) we should have an exponential which depends on t to compensate. Additionally, their $\Delta = O(1)$ in lattice units and their $t < 10$ (in lattice units). In that regime, we might very well have $O(\Delta t \Gamma(t_0)) = O(\Gamma(t_0))$, and it's unclear how one could distinguish the two expressions.

Another concern one might also have is the shift [37] applies to the (A) sector is dependent on t, t_0, so we should be concerned about the error terms generated. However, the shift is simply $e^{-\Sigma t}$, where Σ is

$$\Sigma(t,t_0) = E_n^{eff,B}(t,t_0) - E_m^{eff,A}(t,t_0) \tag{8.21}$$

This shift exponential cancels in the ratio, but we are left with a first order error term from eq. (8.21) which has t, t_0 dependence:

$$O(e^{O(e^{-\Delta t})})$$

(see section 8.1) which gives us first order error terms $\Gamma(t)$ in the exact same manner we obtain $\Gamma(t)$ before. Thus, our expression in eq. (8.20) is unchanged.

8.4 Around the World (ATW) Artifacts

Assume we have a two particle state with energy E. The particles can both travel forward in time giving us $\sim exp(-Et)$, backwards in time giving us $\sim exp(-E(L_t - t))$, or in opposite directions in time. Assuming the particles on average have energies E_1, E_2 where $E = E_1 + E_2$, this gives us a piece

$$e^{-E_1 t} e^{-E_2(L_t - t)} = e^{-(E_1 - E_2)t} e^{-E_2 L_t}$$

which vanishes in the limit $L_t \to \infty$. We thus identify this amplitude as an artifact of the finite size of our time extent L_t. We are generally not able to deduce E_1, E_2 analytically from E, which turns our single state fit to E into two state fits (which we are trying to avoid in general) to $E, \Delta E \equiv E_1 - E_2$. If we have many two

particle states in our spectrum, we end up with many ATW terms with different ΔE time dependence which contaminate our results. Methods for dealing with these artifacts follow.

8.4.1 Matrix Subtraction

We follow the method of [38]. Starting with our GEVP matrix $C(t)$, we create a new matrix $D(t)$ defined as

$$D(t) = \left(e^{\Delta E t} C(t) - e^{\Delta E (t-\delta_t)} C(t-\delta_t)\right) \tag{8.22}$$

The purpose of the exponentials in eq. (8.22) is to make the ATW term constant allowing for removal via subtraction. The spectrum of D is shifted from C by $-\Delta E$, so we add back in ΔE to get the spectrum with the ΔE ATW piece removed. In our studies, we set $\delta_t = 3$ (as in [38]), $\delta_t = 4$, or $\delta_t = 1$ if the need for time more time slices for our fit outweighed the attendant increase in noise. One loses δ_t time slices in this procedure, which can prove very costly on ensembles with small L_t. We usually apply this procedure twice, setting $D(t)$ to be the new $C(t)$ in the second subtraction. This allows us to remove the two leading order ATW terms. For most cases, matrix subtraction suffices to remove the majority of ATW contamination. However, when we examined the smaller physical time extent of the section 3.3 this method proved insufficient.

For an explanation of the effective mass procedure for $D(t)$, see section 9.2.

8.4.2 Vacuum Saturation Subtraction

For every two particle energy difference, we get two around the world terms (the other one is reversing the time propagation of both pions). We can simplify somewhat by supposing that the majority of the amplitude of the ATW terms are non-interacting amplitudes. This reduces the maximum number of unique ATW terms to $2N$ for an $N \times N$ GEVP. Indeed, it is easy to imagine that pions which do not occupy the same time slice will exchange fewer gluons. We thus look at the non-interacting amplitude of $\pi\pi \to \pi\pi$. This is obtained by squaring the gauge configuration average of the single pion correlation function (see section 8.10). We do all the subtraction under a jackknife.

For every topology of the non-interacting D diagram (see eqs. (4.1) and (4.2)), we get a different around the world term to subtract. We can obtain the around the world term to subtract in two (asymptotically) equivalent ways:

ATW.1 We subtract the full non-interacting amplitude. However, we also subtract the wanted cosh part of the amplitude where the pions travel in the same direction in time. We thus find the individual single pion

correlation function late time effective energies E_1, E_2, and, using the same two late time slices, find the single pion amplitudes (A_1, A_2). We then compose

$$A_1 A_2 \left(e^{-(E_1+E_2)t} + e^{-(E_1+E_2)(L_t-t)} \right) \tag{8.23}$$

eq. (8.23) then constitutes the desired portion of the non-interacting amplitude we add to compensate.

ATW.2 We know that the GEVP has two time separations (t, t_0). These separations for the single pion correlation functions give us enough information to get A_1, A_2, E_1, E_2. We then can compose and subtract

$$A_1 A_2 \left(e^{-E_1 L_t} e^{-(E_2-E_1)t} \right) \tag{8.24}$$

$$A_1 A_2 \left(e^{-E_2 L_t} e^{-(E_1-E_2)t} \right) \tag{8.25}$$

ATW.3 We can also use the jackknifed dispersive energies from the pion mass and a single time slice to obtain the amplitude.

While item **ATW.2** might be less noisy since it uses earlier times and is more correlated with the resulting GEVP, we end up with more systematic error. In fact, we can regard item **ATW.2** as the dual limit of item **ATW.1** as the two times used for the effective mass go to t, t_0.

It might be advantageous in certain situations to consider other time slice pairs for item **ATW.1** (e.g., for $t - t_0 \gg 1$ we might choose $t_0, t_0 + 1$).

item **ATW.3** is likely less correlated than with the interacting amplitude than, say, item **ATW.2**. One might obtain smaller errors, however, due to the constant signal to noise ratio of the stationary pion.

8.4.3 Measure the ATW Terms Directly

A study of this method on our data is beyond the scope of the present work, but one might well imagine measuring directly quantities like $\langle \pi | (\pi\pi)_{I=2} | \pi \rangle$.

8.4.4 Extract Via Simultaneous Fit

If one forgoes the GEVP entirely, one can simultaneously fit the correlation functions in $C(t)$. If one has ATW terms in the fit function, one can in principle disentangle the physical states from the ATW terms. One encounters the same challenges as fitting to more than N states in an $N \times N$ GEVP. A study of this method on our data is beyond the scope of the present work.

8.5 Generic Prescription (non-GEVP) for ATW Errors in Matrix Elements

(One might very well object that this is not strictly a GEVP topic. However, we place this section in chapter 8 because of the similarity of its contents to other sections in this chapter.)

Suppose we want to calculate a lattice three point function, but we are concerned that there might be significant around the world contributions. Call the source, sink operators B, A respectively.

$$\begin{aligned}\langle A(t_2)|H(t_1)|B(0)\rangle = &\langle 0|A|A\rangle\langle A|H|B\rangle\langle B|B|0\rangle e^{-E_A t_2}e^{-t_1(E_B-E_A)}+\\ &+\langle A|A|0\rangle\langle 0|H|AB\rangle\langle AB|B|A\rangle e^{-E_A(T-t_2)}e^{-E_{AB}t_1}\\ &+\langle B|A|AB\rangle\langle AB|H|0\rangle\langle 0|B|B\rangle e^{-E_B(T-t_2)}e^{-E_{AB}(t_2-t_1)}\\ &+\langle AB|A|B\rangle\langle B|H|A\rangle\langle A|B|AB\rangle e^{-E_{AB}(T-t_2)}e^{-E_B t_2}e^{-t_1(E_A-E_B)}\\ &+\ldots\end{aligned}$$

Suppose we work in momentum space and average over $\vec{p}, -\vec{p}$. Let's examine the second term and apply crossing symmetry:

$$\begin{aligned}\langle A|A|0\rangle\langle 0|H|AB\rangle\langle AB|B|A\rangle e^{-E_A(T-t_2)}e^{-E_{AB}t_1} &= \langle 0|A|A\rangle^*\langle A(-t_1)|H|B\rangle\langle AB|B|A\rangle e^{-E_A(T-t_2)}e^{-E_{AB}t_1}\\ &= \langle 0|A|A\rangle^*\langle A(T-t_1)|H|B\rangle\langle AB|B|A\rangle e^{-E_A(T-t_2)}e^{-E_{AB}t_1}\\ &= \langle 0|A|A\rangle^*\langle 0|A|A\rangle\langle A|H|B\rangle\langle AB|B|A\rangle e^{-E_A(2T-t_2-t_1)}e^{-E_{AB}t_1}\\ &= |\langle 0|A|A\rangle|^2\langle A|H|B\rangle\langle AB|B|A\rangle e^{-E_A(2T-t_2-t_1)}e^{-E_{AB}t_1}\end{aligned}$$

Similarly, we find for the third term

$$\langle B|A|AB\rangle\langle AB|H|0\rangle\langle 0|B|B\rangle e^{-E_B(T-t_2)}e^{-E_{AB}(t_2-t_1)} = \langle B|A|AB\rangle\langle A|H|B\rangle|\langle 0|B|B\rangle|^2 e^{-E_B(2T-t_2-t_1)}e^{-E_{AB}(t_2-t_1)}$$

Now, inserting a complete set of states twice we find (e.g.)

$$\begin{aligned}
\langle AB(t_1)|B(0)|A(0)\rangle &= \sum_{n,m} \langle A(t_1)|m\rangle \langle m|B(t_1)|B|n\rangle \langle n|A(0)\rangle \\
&= \sum_{n,m} \langle m|B(t_1)|B|n\rangle \langle A(t_1)|m\rangle \langle n|A(0)\rangle \\
&= \langle A(t_1)|A(0)\rangle \sum_n \langle n|B(t_1)|B|n\rangle + \sum_{n\neq m} \langle m|B(t_1)|B|n\rangle \langle A(t_1)|m\rangle \langle n|A(0)\rangle \\
&= \langle A(t_1)|A(0)\rangle \left(\langle B(t_1)|B(0)\rangle + \sum_{n\neq 0} \langle n|B(t_1)|B|n\rangle \right) + \sum_{n\neq m} \langle m|B(t_1)|B|n\rangle \langle A(t_1)|m\rangle \langle n|A(0)\rangle \\
&\equiv \langle A(t_1)|A(0)\rangle \langle B(t_1)|B(0)\rangle + \text{interacting terms}
\end{aligned}$$

Suppose we have two point data for A,B at $t = t_2 - t_1, t_1$ which to a good approximation can be modeled by a single exponential (or can be forced to be so at these time separations given that data). We can then conclude

$$\frac{\langle A(t_2)|H(t_1)|B(0)\rangle}{\langle A(t_2)|A(t_1)\rangle \langle B(t_1)|B(0)\rangle} = \langle A|H|B\rangle \left[1 + \langle A|A\rangle e^{-2E_A(T-t_2)} e^{E_A t_1} + \langle B|B\rangle e^{-2E_B(T-t_1)} e^{E_B t_1} + \text{interacting}\right]$$

Neglecting the interacting pieces, we can divide out the term in brackets to get our improved $\langle A|H|B\rangle$

This note benefited from notes by X. Feng (priv. comm.).

8.6 Positivity

Theorem 2. *The GEVP matrix should be positive semi-definite.*

Proof. Neglecting the finite size of the box, we find

$$\begin{aligned}
C_{ij}(t) &= \left\langle 0|O_i O_j^\dagger|0\right\rangle \\
C_{ij}(t) &= \sum_{n=0}^{\infty} \left\langle 0|O_i(t)|n\rangle \langle n|O_j^\dagger(0)|0\right\rangle \\
C_{ij}(t) &= \sum_{n=0}^{\infty} \psi_{in} e^{-E_n t} \psi_{jn}^* \\
\langle v|C|v\rangle &= \sum_{i=0}^{\infty} \left\langle v|\psi_n\right\rangle e^{-E_n t} \left\langle \psi_n|v\right\rangle \\
&= \sum_{n=0}^{\infty} |u_n|^2 e^{-E_n t} \\
\Rightarrow \langle v|C|v\rangle &\geq 0
\end{aligned} \tag{8.26}$$

We note that adding terms which involve travel through the boundary essentially adds the complex conjugate of the RHS of eq. (8.26), so does not change the conclusion.

□

8.7 Wandering Eigenvalues

When one solves the GEVP, one gets a set of N eigenvalues, which are not necessarily ordered. One can sort these eigenvalues, but the sorting will fail if the errors on some eigenvalue are of the same order as the difference in the mean of that eigenvalue and the mean of its adjacent neighbors. Some samples of one eigenvalue, then, may enter into the collection of samples of a different eigenvalue (causing the sample averages to become biased). These eigenvalues are said to "wander." There are two known ways to deal with this problem.

8.7.1 ϵ Prescription

We take the ith sample of the GEVP matrix $C(t)_i$ and define $C(t)_i \to \left(C(t)_i - \overline{C(t)}\right)\epsilon + \overline{C(t)}$ for some small $\epsilon > 0$ where the overbar indicates an average over index i. We then take the resulting energies processed from the GEVP E_i and apply $E_i \to \frac{1}{\epsilon}\left(E_i - \overline{E}\right) + \overline{E}$. Notionally, we identify $\langle z_i \rangle \equiv \overline{z}$ to the sample average of parameter z with respect to sample index i (which we know is an unbiased estimator of z).

Theorem 3. *This is an unbiased estimator of E to order ϵ. More generally, applying this procedure to any function f of some parameters x^b to yield y will give an unbiased estimator of y.*

Proof. Assume $y \equiv f(x^b)$, that is $\overline{y} = f(\overline{x}^b)$. Then,

$$\begin{aligned} y_i &= \left(f(\epsilon(x_i^b - \overline{x}^b) + \overline{x}^b) - \overline{y}\right)\frac{1}{\epsilon} + \overline{y} \\ &= (x_i^b - \overline{x}^b)\partial_b f(\overline{x}^b) + \overline{y} + O(\epsilon) \\ \Rightarrow \langle y_i \rangle &= \overline{y} + O(\epsilon) \end{aligned} \tag{8.27}$$

□

This proof is analogous to showing the jackknife procedure allows us to linearize the error dependence of y on x. eq. (8.27) is a beginning ingredient in the usual proof of the linear dependence of the errors in y on x and ϵ in the usual setup might be considered a parameter to keep track of the expansion order, rather than something we explicitly modify as in the above procedure. In fact, if the samples are jackknife samples, the expansion can be considered $O\left(\frac{\epsilon}{N}\right)$ since the bias in $\overline{y}$ compared to $\langle y \rangle$ (the true average) is $O\left(\frac{1}{N}\right)$. Thus, using a smaller ϵ will allow for more linearity in the small N limit.

The above method was first invented and applied (by C. Lehner) to deal with wandering eigenvalues in the distillation $I = 1$ $\pi\pi$ data. We later saw the same phenomenon appear in the $I = 1$ A2A data (as expected).

8.7.2 Tracking Via Eigenvectors

Briefly, one might also consider taking the eigenvectors as indicators of when an eigenvalue has wandered. The validity of this method rests on the assumption that the eigenvectors have smaller variance and so do not wander. The procedure is as follows:

1. We find the average eigenvectors $\overline{v}^a$ ($0 <= a <= N-1$ is the operator index for an $N \times N$ GEVP).

2. For the ith sample eigenvectors v_i^b, form $v_i^b \cdot \overline{v}^a$. We know that if the variance in the eigenvectors is small enough that the eigenvectors do not wander, $v_i^a = max(v_i^b \cdot \overline{v}^a)$ maximized over index b.

3. Once we find the one-to-one mapping of index $b \rightarrow a$, we can use the same mapping to identify eigenvalues from sample i with their averages.

This method has not yet been implemented to the author's knowledge. The ϵ prescription (see section 8.7.1) is slightly easier to implement, and numerical evidence w.r.t. the size of the errors in the GEVP vs. the errors in the resulting energies suggests it solves the problem (for sample sizes so far seen; These sample sizes do display the property that increasing the number of samples decreases the errors, also known as the Gaussian regime).

However, these methods are not mutually exclusive, so it might be interesting to consider applying both (especially when the sample size is small, and/or if one is skeptical of the results from the ϵ prescription).

8.8 Deleting GEVP Operators

It is sometimes apparent that truncating the operator basis gives smaller errors overall. We thus might conclude that one operator is only contributing noise. It is an open problem to describe precisely when/how to remove operators. We might remark, however, that truncating the operator basis for only a subset of the time slices makes it impossible to identify the systematic errors due to excited states at the boundary between different operator sets. This might still be worthwhile if the fits are to a constant, however.

Given the finite number of possibilities to test, looping over different operator deletions might be the best method to determine which operator set to use. However, this can of course miss an important state.

8.9 Minimal Estimate of Excited States Contamination

Theorem 4. *Suppose we have two point data for an arithmetic sequence of at least three time slices we can label w.l.o.g. $t > t_0 > t_0'$ (where $\alpha \equiv t - t_0 = t_0 - t_0'$). Then we can estimate the size of the first order excited state contamination $O(e^{-\Delta t})$ (see section 8.1) up to a (first order) correction in Δ of size $O(\frac{1}{t_0'} e^{-\Delta\alpha})$.*

Proof. With the given data we can form GEVP time pairs (t, t_0) and (t_0, t_0'). . When we take the log of the eigenvalues, we get (from section 8.1)

$$\log(\lambda(t, t_0)) = E_n \alpha + O(e^{-\Delta t_0})$$
$$\log(\lambda(t_0, t_0')) = E_n \alpha + O(e^{-\Delta t_0'})$$

Subtracting, we find

$$\begin{aligned} -\frac{1}{t_0'} \log(\log(\lambda(t_0, t_0')) - \log(\lambda(t, t_0))) &\approx -\frac{1}{t_0'} \log(O(e^{-\Delta t_0}) + O(e^{-\Delta t_0'})) \\ &= \Delta - O(\frac{1}{t_0'} e^{-\Delta\alpha}) \end{aligned}$$

□

This method is minimal in the sense that it uses the fewest time slices possible. We thus limit the statistical noise in this estimate. This method so far has been tested using small GEVP matrices of fake data, where it was found that if the coefficient of $e^{-\Delta t_0}$ is not close to 1 in absolute value the convergence rate might be too slow to be useful. However, if one has obtains a fit where the coefficients are $O(1)$, one can use this method to perform a quick verification of the excited state contamination.

8.10 Pion Ratio Method

This method was originally investigated by X. Feng [39] as a way to reduce noise in $I = 2$ $\pi\pi$ scattering and subsequently explored by D. Murphy [40]. The author learned initially of the early form of this method from D. Murphy and had the idea to extend it to the $I = 2$ excited states (the author was ultimately unsuccessful due to not appreciating how delicate the correlations must be to get smaller errors; the method was tried on the eigenvalues of the GEVP which was very naive in retrospect). However, the problem was eventually solved by C. Lehner a short time after he began studying the $I = 2$ spectrum. We outline his method in this section.

Roughly, if we square the averaged pion correlation function, we will get a two pion correlation function with energy of a two pion non-interacting state (which is not a physical quantity, but something well-defined

nevertheless). If we add pion energies obtained from a fitted pion mass and Einstein's dispersion relation $\sqrt{m^2+p^2}$, we will also get the energy of a two pion non-interacting state. If we take the difference between these two energies, we will get something which is statistically 0 in the continuum limit (the latter distinction is important since the dispersion relation might give different results than the pion correlation function squared due to lattice spacing artifacts. Assuming we are subtracting the square and adding in the dispersive energies, we cancel some of this artifact.). We then add this to the interacting energies (which we are allowed to do since we are effectively adding 0).

It turns out that if one replicates the time folding (see section E.1 for definition) and $t_{\text{ src}}$ averaging (due to time translation symmetry) and topologies of the interacting diagrams exactly in the square of the pion correlation function, one can reduce the statistical noise by a factor of 10. However, one must be careful to jackknife the pion mass estimates because the resulting errors should be (roughly) bounded by the errors on the pion mass. One should obtain a Pearson / Spearman ([41]) R coefficient > 0.9 with high confidence between the interacting and non-interacting energies (from the square). This result generalizes back down to the correlation function level.

We can obtain this result by solving (for the two pion non-interacting state) a set of 1×1 GEVP's. We know the off-diagonal terms which appear in the interacting GEVP are all zero due to momentum conservation of each separate pion.

The pion ratio method has the nice feature of automatically correcting discretization errors, because we subtract a fit non-interacting energy and add in something derived from the (continuum) dispersion relation. Thus, if we had for instance a lattice dispersion relation from $\hat{p}$ (see section 10.5.2), this difference would be statistically non-zero and would correct this violation of the continuum dispersion relation. It also tends to cause earlier plateaus since the pion mass usually has less excited state contamination for earlier times than single pions with momentum.

Unfortunately, this result does not give as great a noise reduction in $I = 0,1$ likely due to our inability to decompose $\overline{\psi}\psi, \overline{\psi}\gamma_\mu\psi$ into a similar product of two operators. However, this method does motivate the search for as yet undiscovered quantities which are also correlated zeros to some quantity of interest. It also shows us that simply because a quantity is not physically meaningful is not necessarily a reason not to consider calculating it as a way to control various systematic or statistical errors.

Much of this code was later adopted for use in vacuum saturation subtraction (see section 8.4.2).

8.11 GEVP Derivative

If we have $\lambda = e^{-Et}$ we have several ways to extract E. We present two effective mass methods to extract E (other options include extraction via fit or use either of the strategies from sections 8.2.2 and 8.2.3 if the GEVP is special in the sense specified there; it is an open problem whether there are other strategies):

$$E = -\frac{1}{t}\log(\lambda) \tag{8.28}$$

$$E = \log\left(\frac{\lambda(t,t_0)}{\lambda(t+1,t_0)}\right) \tag{8.29}$$

It turns out eq. (8.29) gives oscillating effective masses which tend to verge on statistical disagreement at adjacent time slices. However, if one examines the form of section 8.1 or section 8.2, we can see eq. (8.29) eliminates terms which are dependent only on t_0. This is how the leading order systematic error in section 8.1 is $O(e^{-\Delta t})$ rather than $O(e^{-\Delta t_0})$. In general, we do not use section 8.11 as the oscillation outweighs gains in systematic error elimination (which we handle via fits; see section 9.7).

8.12 GEVP Eigenvectors

Besides their (hypothetical) use in section 8.7.1, we might suppose GEVP eigenvectors are only incidental to solution of the GEVP. This is mostly true. The eigenvectors tend to be noisier, have systematic errors which decay more slowly than the energies (see eq. 2.23 of [35]), and are harder to resolve over time due to the larger number of fit parameters needed. However, since we get them for free in the course of solving the GEVP, we are still motivated to suggest possible uses (none of which has been extensively explored by the author).

8.12.1 Transform to a Scalar Problem

The eigenvectors could be hard to fit, so we should think about collapsing their information into a scalar which we can fit. We also know the eigenvectors might, in many cases, be almost orthogonal so we can form this scalar via the norm of a cross product among the vectors. This procedure amplifies the systematic error we want to fit. We also know the time dependence is only in the systematic error, so finding this cross product at fixed t but different t_0 allows us to perform a GEVP derivative which eliminates everything but a term like $\Gamma(t)$ (see section 8.3.1). We can then find the effective mass of that term and use this to estimate E_{N+1} as well as the error decay rate in our problem.

8.12.2 Use Eigenvectors From a Different Times

Starting from our definitions of A, B in section 8.1, we have

$$\begin{aligned} A|v\rangle &= \lambda B|v\rangle \\ \Rightarrow \lambda &= \frac{\langle v|A|v\rangle}{\langle v|B|v\rangle} \end{aligned} \tag{8.30}$$

This implies that any eigenvector $v' \approx v$ can give us an approximation for λ via the ratio in eq. (8.30). Furthermore, from section 8.2, we know the systematic error of the eigenvectors enters at second order in our expansion parameter. Thus, even though it decays via a t_0 dependence instead of via a t dependence, we know its contribution to λ is t dependent (assuming the t, t_0 conditions of section 8.1). Thus, if the eigenvector signal to noise decays quicker (empirical observation) than the signal to noise of the energies, we can simply ask the question whether using eigenvectors from earlier times will lower our overall error by trading systematic error for statistical.

8.12.3 Independence of Eigenvector Problem From GEVP

In principle, we should be able to find the time independent eigenvectors $|v^0\rangle$ via a fit, freeze the result, and then use them to find eq. (8.30). However, this may prove difficult in the low statistics regime. The number of fit parameters needed is $\binom{N}{2}$ for an $N \times N$ GEVP since we have a normalized basis.

8.13 GEVP (2-point) Best Practices

1. Use $t - t_0 = const, t \geq 2t_0$ (see section 8.1).

2. If one can identify a way to extract non-interacting energies separate from the dispersion relation as in section 8.10, then add the difference between this value and that derived from the dispersion relation to the interacting energies. If it reduces noise, use this value for the interacting energy.

3. Check for positivity (section 8.6), or that the LHS and RHS GEVP matrices have the same sign (both pos-def or both neg-def).

4. Remove ATW Terms. If one doesn't measure the terms explicitly, then in the case of the center of mass frame use matrix subtraction. Otherwise, use vacuum saturation subtraction (section 8.4.2).

5. Use the ϵ prescription (section 8.7.1) (especially for $I = 1$).

6. Fit using eq. (9.3) unless the GEVP is special in the sense of either sections 8.2.2 and 8.2.3 (solvable or diagonal in contamination). One might not be sure if section 8.2.2 is applicable, so one should always check with a fit to eq. (9.3). We might identify section 8.2.2 if a single Δ does not suffice for all the eigenvalues (each might be better modeled with a separate Δ_n). However, use of sections 8.2.2 and 8.2.3 are untested.

7. Don't use a GEVP derivative section 8.11 (as it usually adds noise). It is likely that using more than two time slices to determine a single effective mass point will add more statistical error than systematic error it subtracts, but this has not been fully tested.

8. Use known results from phenomenology crossed with Lüscher-type curves to set up the appropriate systematic error ansatz. (Precise method is an open problem).

Chapter 9

Fitting

9.1 Correlated Fitting Under the Jackknife

Briefly, we fit under a jackknife (also known as doing a double jackknife). We do a separate fit for each single jackknife block. Since we are performing correlated fits, we must separately estimate the covariance matrix for each fit (which is to a set of data with one sample left out). We thus leave out another sample to estimate the covariance matrix (so it is the result of two jackknifes).

9.1.1 Correlation Basis Vs. Covariance Matrix

The correlation matrix can be obtained from a covariance matrix by dividing the off-diagonal elements indexed by i, j with the product of the square roots of the i, i element and the j, j element. One can show χ^2 obtained from the correlation matrix or from the covariance matrix should be equivalent if one multiplies the difference vectors which sandwich the covariance matrix by the same factors. In fact, these basis transformation factors are simply the inverse of the diagonal covariance matrix ($Diag(C_{\text{cov}}) \equiv D$). Thus,

$$\begin{aligned} C_{\text{corr}}^{-1} &= (D^{-1} C_{\text{cov}} D^{-1})^{-1} \\ &= D C^{-1} D \\ \Rightarrow C_{\text{cov}}^{-1} &= D^{-1} C_{\text{corr}}^{-1} D^{-1} \end{aligned}$$

Our procedure is thus to rotate to the correlation matrix basis to invert the covariance matrix, then we rotate back to the covariance matrix basis. There is apparently some lore which indicates there are certain

situations when this might be a numerically safer procedure. However, the author has so far not seen this to make a difference but applies it nonetheless.

9.2 Effective Mass Procedure

The basic idea of effective mass processing is to force a fit to extract the parameters of interest by reducing the degrees of freedom (dof) to 0. Most often, these parameters are energies, which we refer to as effective masses or effective energies. In most cases, one performs this forced fit on every time slice using extra time slices to reduce the dof. We refer to this set as the minimal subset. These extra time slices have a fixed relationship to the original time slice. Afterwards, one has an estimate of the fit parameters for every time slice. One may select a particular time slice for the final estimate (chosen before the processing to avoid bias), or fit to the effective masses.

There are, broadly, four classes of effective mass methods.

EF.1 Exact function inverse. We find an exact inverse function (like arcCosh), which we can apply to our minimal subset which yields our parameter(s) of interest.

EF.2 We numerically solve a system of equations such that the number of time slices in our minimal subset is equal to the number of fit parameters.

EF.3 We perform a global fit to $r(t)$ (see eq. (9.1)) directly.

EF.4 We minimize some cost function $g(h(t) - r(t))$ where g is the cost function, $r(t)$ is the processed ratio (some function of our minimal subset), and $h(t)$ is our fit function. This cost function should have only one solution. This is the traditional effective mass method.

We have found that, in general, item **EF.1** is noisy due to the large number of time slices needed (which may be less than perfectly correlated), item **EF.2** is numerically unstable, and item **EF.3** is a computationally costly version of item **EF.4**.

We therefore ended up using item **EF.4** exclusively. Our cost function was $g(t) = (h(t) - r(t))^2$. In some cases, this yields a positive and negative solution with similar costs. We chose the physically more meaningful (positive) solution as long as it does not exceed some fixed cost (set arbitrarily at 10^{-12}).

In the remaining portion of this section, we define $r(t)$ (see section 9.2.1) and $h(t)$ (see section 9.2.2).

9.2.1 Data Ratios

For single correlation functions, we define

$$r(t) = \log\left(\frac{C(t)}{C(t+1)}\right) \tag{9.1}$$

We briefly studied omitting the log, but it seemed not to make a substantial difference. In the case of a state with a single exponential with energy E, $E = r(t)$, so using the log makes $r(t)$ more physically meaningful.

For the GEVP, we minimize a different cost function for each eigenvalue of the GEVP matrix. Thus, each $r(t)$ is equal to the log of each of the eigenvalues if we are not taking a GEVP derivative (see section 8.11). If we are taking a GEVP derivative, we form eq. (9.1) (for each dimension). This derivative eliminates first order t_0 systematic error due to excited states, and is explained extensively in section 8.1. Otherwise, each $r(t)$ is just the log of each of the eigenvalues.

9.2.2 Fit Ratios

The forced fit function $h(t)$ can be quite complicated depending on how much processing is done to the correlation function(s). For single correlation functions, we use

$$\begin{aligned} cc(t) &\equiv e^{-Et} + e^{-E(L_t - t)} \\ h(t) &= \log\left(\frac{cc(t)}{cc(t+1)}\right) \end{aligned} \tag{9.2}$$

GEVP without matrix subtraction (eq. (8.22)), and a GEVP derivative (eq. (8.29)) we use eq. (9.2) for each dimension. Without a GEVP derivative, we use

$$h(t) = \log\left(\frac{cc(t)}{cc(t_0)}\right)$$

for a GEVP LHS, RHS time pair t, t_0. With matrix subtraction (eq. (8.22)), we substitute

$$cc(t) \to cc(t) - cc(t - \delta_t)$$

For two matrix subtractions with (in general) different δ_t ($\equiv dt_1, dt_2$)

$$cc(t) \rightarrow (cc(t) - cc(t - dt_1)) - (cc(t - dt_2) - cc(t - dt_1 - dt_2))$$

9.3 Fit Ranges

9.3.1 Introduction to Fit Range Averaging

This section is intended to explain the idea of fit range averaging, that is, the technique of fitting to different subsets of data and then averaging the extracted fit parameters.

Why average over fit ranges? This is a good question, and one that should be answered before applying this technique. In one case, we fit to a constant and then find a function of the small deviations of this constant from a known value. Importantly, the size of statistical fluctuations in the constant mean measuring (or fitting to) the function of the deviations (phase shifts) may never be numerically stable enough to be viable. Thus, the general statement is that the fitting to different subsets of the data does not place strong constraints on our model or on the model parameters. This leads to multiple different models (or model parameters) which fit different subsets of the data, and these may be incompatible in the low statistics regime.

We could simply average over the fit points (since we are fitting to a constant), taking into account the correlations to propagate the uncertainty. In general, this procedure is to find the minimal sized fit range which force given values for model parameters, and then averaging over these. However, this discards information about fit quality lending equal weights to fits with unequal systematic errors (which the fit quality is somewhat sensitive to). Also, averaging over fit points discards information about the correctness of fitting to the given model in the first place. We may have, for example, unavoidable excited state contamination in some of the fits ranges. If we average the fits (or take the median), we are able to weight contributions from this systematic error less.

The other limit is to use only the single maximum fit range. This usually only is viable in the high statistics limit, which may be expensive to reach (this is definitely true in our case).

Even if the fit succeeds, some small eigenvalues of the covariance matrix may cause the fit to be less useful for estimating parameters, so sampling different fit ranges allows us to avoid having our statistical error swamped by a single systematically poor fit (see section 9.4). Already, the unweighted average samples data more often if it gives compatible fit results, which means it samples based on the information we have about the correlations. If we then weight the fit ranges based on the goodness of fit, we essentially preserve this information out to the level of the final average (see section 9.4.1). We therefore identify section 9.4.1 as the

version of this technique most representative of the population average (which we obtain with good statistics and good control over the systematics).

Finally, this technique is useful because it removes human bias in selecting a fit range. One might respond that fitting to a subset of data is something that one could simply exclude from the analysis in the first place (as it at least decreasing precision). In a long running calculation, one might therefore simply wait until enough data accumulates before performing the analysis. This, of course, would give unbiased results. However, it is probably very often the case that one would instead like to perform a low statistics calculation, and so taking a subset of the data (say, by skipping every other data point) may increase the goodness of fit. In the high statistics limit, fit range averaging, fitting to a single fit range, and averaging over subsets of fit ranges must all give compatible results. Thus, the main reason to use fit range averaging is if one has limited statistics and the desire to preserve information about correlations out to the level of the model parameters one would like to extract. However, in the case of lattice QCD data, we often find that the data is highly correlated, so averaging does not usually improve statistics.

9.4 Unweighted Averaging

We start by assuming that our model will give good fits, and will not overfit. Thus, we have need for a well-chosen fit "window". Fit windows are defined via a pair of time slices, $t_{\text{ min}}, t_{\text{ max}}$ such that all subsets included in the fit have $t_{\text{src}} >= t_{\text{ min}}, t_{\text{snk}} <= t_{\text{ max}}$. One can automate this selection to some extent by adding systematic error estimators to the model (if underfitting) and subtracting data with large noise (relative size of error bars, e.g.) (if overfitting), but this is another topic. In any case, one should vary this window to make sure it is well chosen, and partition the set of fit ranges to verify consistency (and stability) of the average. Our window amounts to a fit range over which our model (and its free parameters) should not vary much in the large statistics limit.

Once our fit window is found, any fit selection (subset of data within that window) will give us an unbiased estimator for our model parameters. However, in the low statistics limit, we know the parameters may vary over fit selections. If we do an uncorrelated average over N fit ranges each of which is fit to the same underlying set of M jackknife samples to obtain some parameter X, we obtain $\overline{X}$:

$$\overline{X} = \frac{1}{N}\sum_i^N X_i$$
$$\sigma_{\overline{X}} = \sqrt{\sum_{i,j}\sigma_{ij}\frac{\partial\overline{X}}{\partial X_i}\frac{\partial\overline{X}}{\partial X_i}}$$
$$= \frac{1}{N}\sqrt{\sum_{i,j}\sigma_{ij}}$$
$$\sigma_{ij} = \frac{M-1}{M}\sum_k^M (X_{ik}-\overline{X_i})(X_{jk}-\overline{X_j})$$

9.4.1 Weighted Averaging

Suppose we have instead weights $w_i \equiv \frac{1}{(\chi^2/dof)_i}$ for a weighted average. The fits which are better (smaller χ^2/dof) get weighted more. This may work, but has the troubling property that the weights are not normalized in a meaningful way. We instead should weight by the p-value (1 minus the cumulative probability distribution of χ^2 for a given number of degrees of freedom; it measures the probability that we would obtain this large of a χ^2 if our data and model were consistent) of obtaining a given χ^2/dof. Suppose, therefore, $w_i = (\text{p-value}(\chi^2/dof))$.

We thus redefine our $\widetilde{X}_i \equiv \frac{X_i w_i}{\sum_j w_j}$.

$$\overline{X} = \frac{\sum_i^N X_i w_i}{\sum_i w_i}$$
$$= \sum_i^N \widetilde{X}_i$$

Since propatation of uncertainty depends only on partial derivatives, the rest of the analysis from section 9.4 follows as before.

9.4.2 Weighted Median

If the frequency plot (histogram) of the results has outliers with non-negligible weight, we might instead consider taking a weighted median. This procedure is much simpler, so will not be explained here.

9.4.3 Fit Range Averaging, Example Results

We document here histograms of $I = 2,\ p11,\ \pi\pi \to \pi\pi$ averaging over $O(1000)$ fit ranges on 155 (zmobius only) configs of the $24ID$ ensemble. The fit window is chosen to be $t = 8-14$, inclusive, which generates $O(10000)$

fit ranges to consider. We compose the fit ranges at random, allowing for duplicates only when the loop is divided over MPI processes. These duplicates are later removed when compiling results. We perform a fit range average weighted by p-value. We are thus left with what turned out to be 997 unique fit ranges. We see good agreement with the example fit results, although the long tails of some of the distributions may mean our requirement that the p-value of the fit ranges we average be > 0.1 is too loose.

Clarification update: The example fit is chosen to minimize the maximum difference between the energy central values of this fit and the fit range average. Thus the tables mainly show the compatibility between the two methods of estimating the statistical errors.

We perform a standard jackknife fit to an example fit range. We find

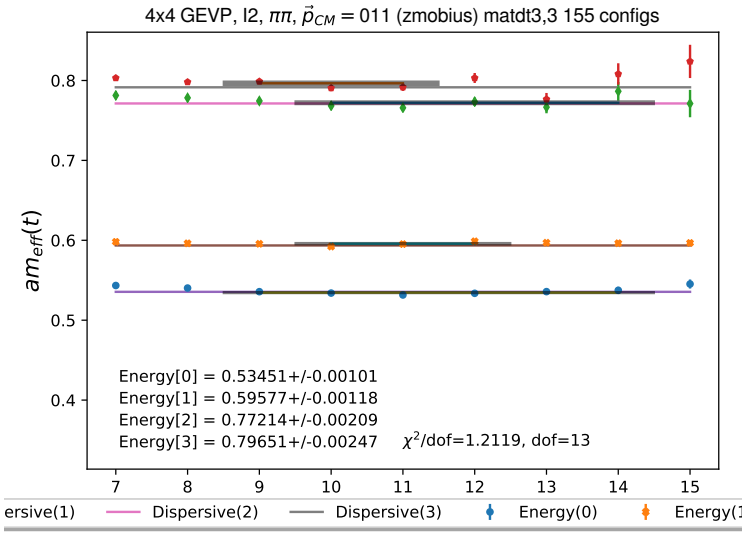


Figure 9.1: Example Fit

State num.	Energy (Example fit, lattice units)	Energy (Fit Range Average)
0	0.5345(10)	0.5347(11)
1	0.5958(12)	0.5958(9)
2	0.7721(21)	0.7720(21)
3	0.7965(25)	0.7966(21)

Table 9.1: Energies

State num.	Phase Shift (Example fit, degrees)	Phase Shift (Fit Range Average)
0	0.42(1.21)	0.22(1.36)
1	-6.6(2.2)	-6.65(1.74)
2	-8.6(6.7)	-9.3(7.0)
3	-14.2(3.9)	-14.3(3.3)

Table 9.2: Phase Shifts

9.4.3.1 Phase Shifts Frequencies

Due to the fact that the error wasn't included in the previous run, these plots are of a different set of 999 fit ranges. However, the results are compatible ($<< 1\sigma$).

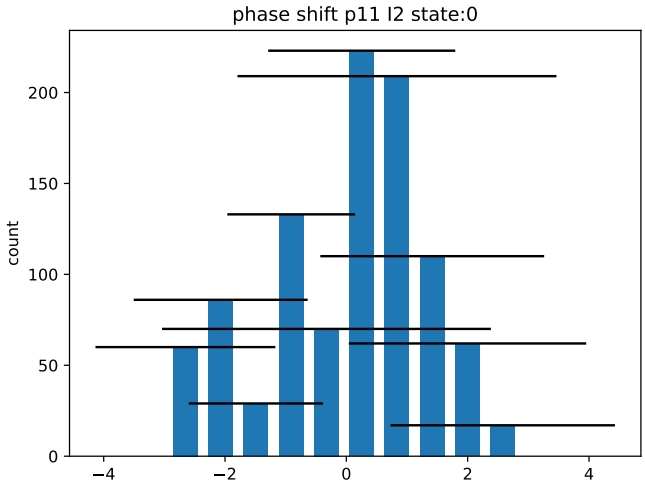


Figure 9.2: ground state phase distribution

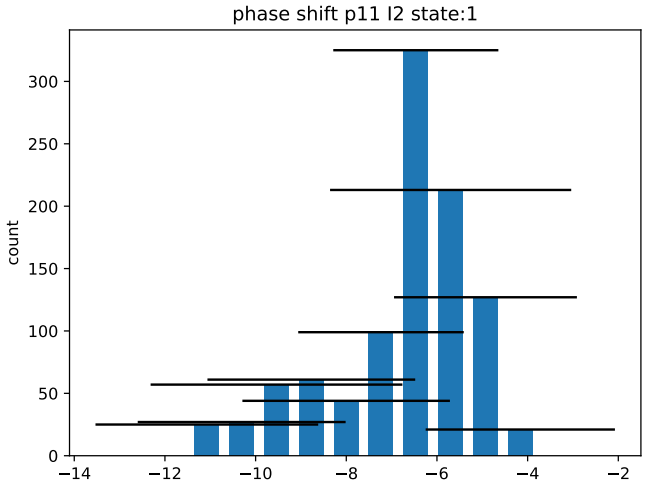


Figure 9.3: first excited state phase distribution

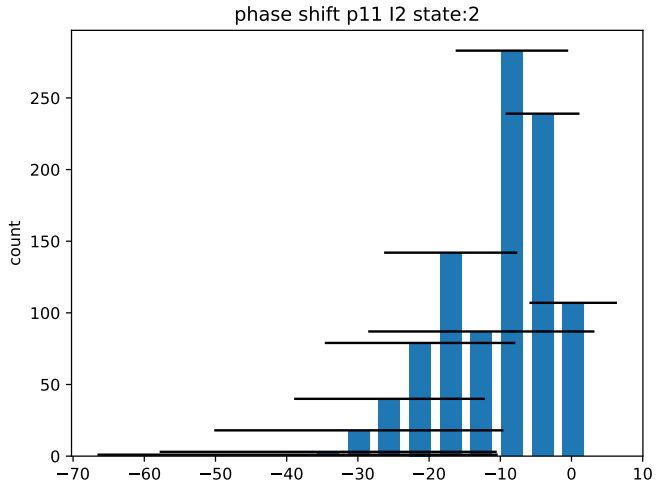


Figure 9.4: second excited state phase distribution

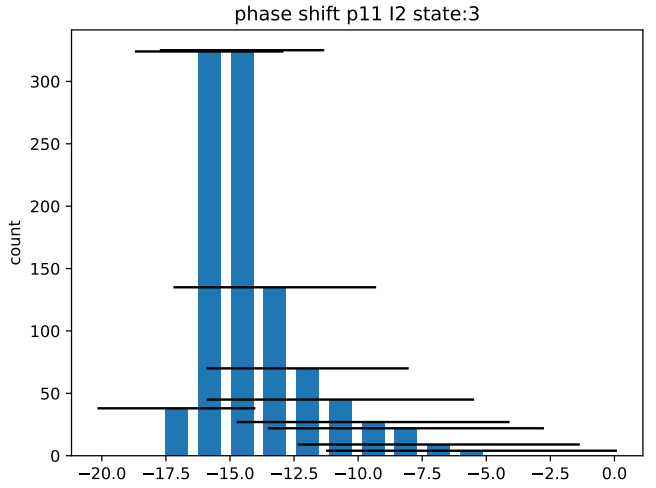


Figure 9.5: third excited state phase distribution

9.4.3.2 Energy Frequencies

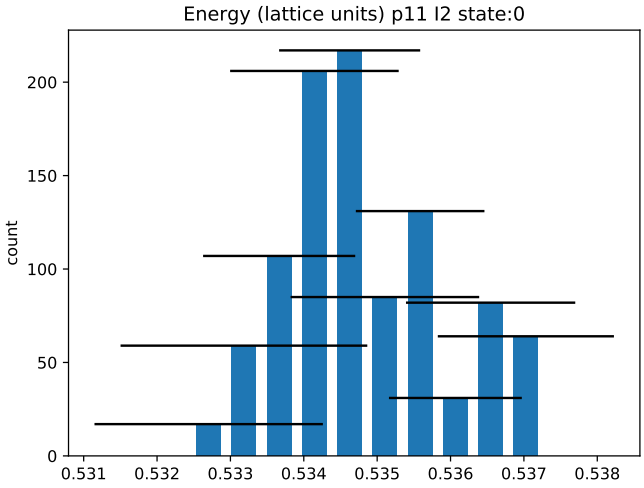


Figure 9.6: ground state energy distribution

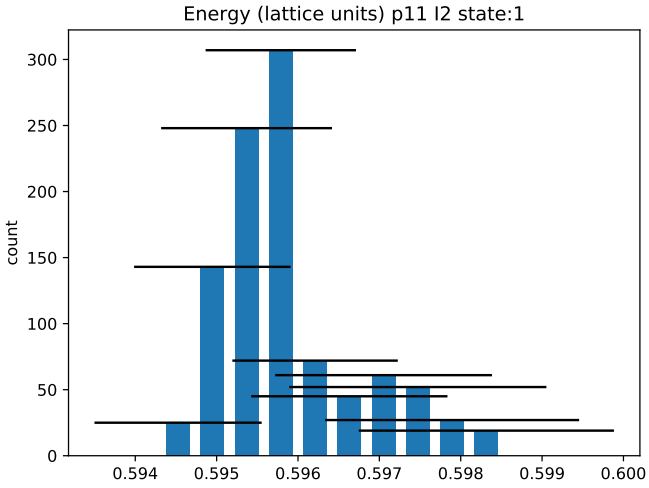


Figure 9.7: first excited state energy distribution

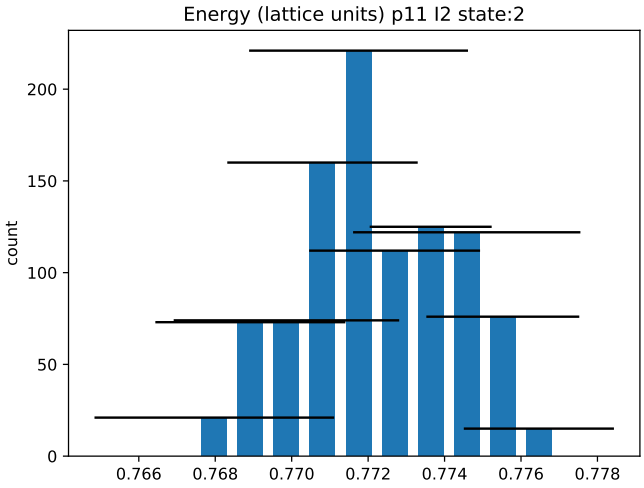


Figure 9.8: second excited state energy distribution

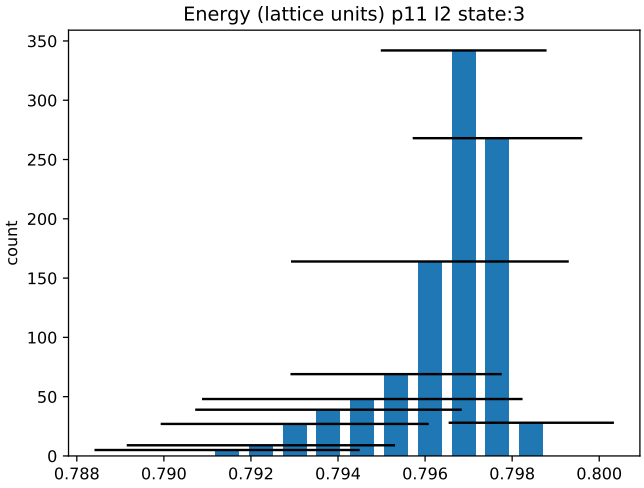


Figure 9.9: third excited state energy distribution

9.4.4 Consistency Conditions, Conclusions

We have defined some tools one might use to study the fit range dependence of our parameter extraction. We have given no reason why one should choose one estimate over another, but if all are unbiased estimators they must all be consistent with each other. In fact, every estimate from every subset of data should be consistent with the averages/medians and with each other. Due to correlations, error bars might be less reliable (especially in the low statistics regime), so this consistency condition might prove difficult to realize in practice. The answer which is forced on us when all estimators are consistent, however, is the estimate with

the smallest errors.

Indeed, for fits to single pion correlation functions, our procedure is to fix t_{max} at a value consistent with our error bar cuts and increase t_{min} until all estimates are consistent with each other. We then pick the estimate with the smallest errors. Usually, this is the effective mass of t_{min}. For our GEVP fits, we usually loop until we get a fit range which gives a good fit, and then we try to extend each dimension in time until the p-value falls below the cutoff. Especially for fits which are not a simple fit to a constant, most fit ranges do not give usable estimates. This paucity eliminates the need to find some consensus answer across fit ranges.

One open problem is to estimate how likely it is that our data is not consistent with the model (e.g. not consistent with the usual cosh ansatz from quantum mechanics) given some fit ranges which give good fits (with acceptable p-values) and some fit ranges which give bad fits (and so are eliminated from our estimates). The reason for a fit range being eliminated is not limited to a low p-value. For instance, if $E < 0$ we mark the fit as bad. Other fit ranges are eliminated because the minimizer does not converge. A question one might have is: how probable are such things? There exists literature on how to deal with multiple hypotheses on the same data set (see, e.g., [42]). However, finding out exactly what this probability (or probabilities) could be seems quite difficult and is beyond the scope of the present work. Presumably there are usually other, better heuristics/checks for finding model-data inconsistencies (for example, some symmetry condition between different correlation functions).

9.5 Analysis/Production Code

Much of the analysis code[15] uses the Numpy library[43]. The plotting routine uses Matplotlib[44]. Processing of hdf5 files is done via H5py[45]. Some check code uses Sympy[46]. The minimizer routines come from Scipy.optimize[47]. The significant figures are handled via gvar [48]. Numeric differentiation is handled via Numdifftools[49]. We also use the minimizer [50] based on [51].

9.6 Cuts, Analysis Choices

9.6.1 Processing and Projections

Due to the time translation invariance of QCD, we are able to average over source operator locations. We refer to the time slice where the source operator is created as t_{src}. We must average over t_{src} as well as fold in time ($f(t) \to \frac{1}{2}\left(f(t) + f(T - t - 2t_{\text{sep}})\right)$). The most important options, in addition to the usual projections onto isospin and angular momentum, (as well as jackknifing and super-jackknifing) are the options for composing

the disconnected diagrams. per section 6.4, we don't take the real part of the bubbles; we do subtraction on moving bubbles, and we don't average the bubbles over t when forming the subtraction.

9.6.2 Fitting

1. We establish an arbitrary cutoff on our p-values of 0.1. Below this value, we say the model does not fit the data and exclude the minimized parameters from our analysis. Importantly, we only apply this cut to the average p-value as we allow for fluctuations about 0.1 in the course of the jackknife loop. We also throw out all fits which are overfit (average $\chi^2/dof < 1$). We throw out fits where the minimizer of χ^2 does not converge. We throw out fits which have a singular covariance matrix anywhere. We throw out fits with imaginary relativistic γ.

2. We also throw out fits if the jackknife distribution exhibits large fluctuations. That is, we examine the first two fits in the part of the superjackknife (see section 2.6) blocks which have fluctuations due to sloppy configurations (thus the fluctuations ought to be smaller so the error on the error is also reduced). If one fit succeeds and the other fails (due to χ^2 being too large or small), we examine the difference in χ^2. If this difference is approximately larger than 5 (an arbitrary amount) standard deviations from 0, we throw out the fit.

3. In these first two sloppy fits, we also apply a cut restricting these two sample χ^2 to either be in the acceptable range for the average, or within 5 standard deviations of the acceptable range. We apply empirically derived corrections when the degrees of freedom are small (and the higher moments of the χ^2 distribution are not negligible).

4. If we have an overflow error anywhere (usually in the minimizer when it guesses too high), we throw out the fit. We throw out fits which have degrees of freedom less than 1. If anything goes wrong in the calculation of the Lüscher zeta function, we throw out the fit. We restrict our fits to have a minimum number of time slices (>= 3). If the error on any of the $I = 2$ phase shifts is > 20 degrees we throw out the fit.

5. We use the error bars from the average covariance matrix. We throw out time slices (for a given operator) if the error bar is $> 20\%$ of the average (usually the latest times). We do this because these will add 1 to χ^2/dof for a wide range of model parameters (they do not constrain the model much). Thus, if we have only a few data points with small errors, their large contributions to χ^2 will be averaged away when dividing by degrees of freedom.

6. We generally use the scipy.optimize.minimize ([47]) minimizer L-BFGS-B to minimize χ^2, but we use the simplex minimizer Nelder-Mead to tune our initial guesses. However, for fits to a constant (effective mass fits), we always start with an initial energy guess of 0.5.

7. We apply section 8.7.1 to our GEVP analysis with $\epsilon = 10^{-4}$ unless we are using the pion ratio method (see section 8.10), where it is not needed (and does not work for some reason).

8. See section 10.5 for our method to correct some of the lattice spacing error by measuring the violation of the continuum dispersion relation.

9. The fit range loop grows exponentially in size with the size of the window. We currently give it a max iteration count (set to 100 at the moment) (below which it runs through all possiblities). If it is beyond this loop count, we randomly sample from the fit ranges until we find > 15 acceptable.

9.7 Systematic Error (Ansatz)

We know from section 8.1 that our effective energies are

$$\begin{aligned} E_n^{eff} &= E_n + O(e^{-\Delta t'}) \\ &= E_n + a e^{-\Delta t'} \\ \Delta &= E_{N+1} - E_n \end{aligned} \tag{9.3}$$

We know that if we form $\frac{\partial E_n^{eff}}{\partial t}$, $t' = t$. If we don't take this GEVP derivative (see section 8.11) and $t - t_0 \neq const$, $t' = t_0$. However, empirically we find that taking the GEVP derivative gives more noise (and oscillating effective masses). This effect appears worse for 32^3. One hypothesis (unconfirmed) is that correlations between adjacent time slices cause the oscillations and the correlations make the error bars artificially smaller. This would explain why the effect is worse for the finer lattice.

Thus, assuming we have some procedure to get the effective energies (usually diagonalizing the GEVP and taking a derivative of the log of the eigenvalues w.r.t. t), can then fit to eq. (9.3). This fit has $2N + 1$ unknowns where N is the dimension of the GEVP.

Chapter 10

$\pi\pi$ Scattering Results

We present various results of physical point $\pi\pi$ scattering, culminating in calculation of phase shifts.

10.1 Measuring Discretization Error on the 24^3

We measure discretization errors via the violation of Einstein's dispersion relation $E^2 = p^2 + m^2$ (which we also call the continuum dispersion relation) on the ensemble detailed in section 3.2.

10.2 Single Pion Energies (zMobius)

P	E (lattice units)	$\sqrt{m_\pi^2 + P^2}$	$\Delta E * 10^5$	stat.
m_π	0.13961(20)	-	-	-
$p1$	0.29612(32)	0.29670(9)	58(33)	1.8 σ
$p11$	0.39431(44)	0.39568(7)	137(45)	3.0 σ
$p111$	0.47223(92)	0.47445(6)	222(92)	2.4 σ

10.3 Lattice Dispersion Relations

This dispersion relation

$$P \to \hat{P} \equiv 2\sin\left(\frac{2\pi n}{24 * 2}\right) \tag{10.1}$$

P	E (lattice units)	$\sqrt{m_\pi^2 + \hat{P}^2}$	$\Delta E * 10^5$	stat.
m_π	0.13961(20)	-	-	-
$p1$	0.29612(32)	0.29604(9)	-9(33)	<1 σ
$p11$	0.39431(44)	0.39470(7)	39(45)	<1 σ
$p111$	0.47223(92)	0.47322(6)	99(92)	1.1 σ

eq. 32 from [52]:

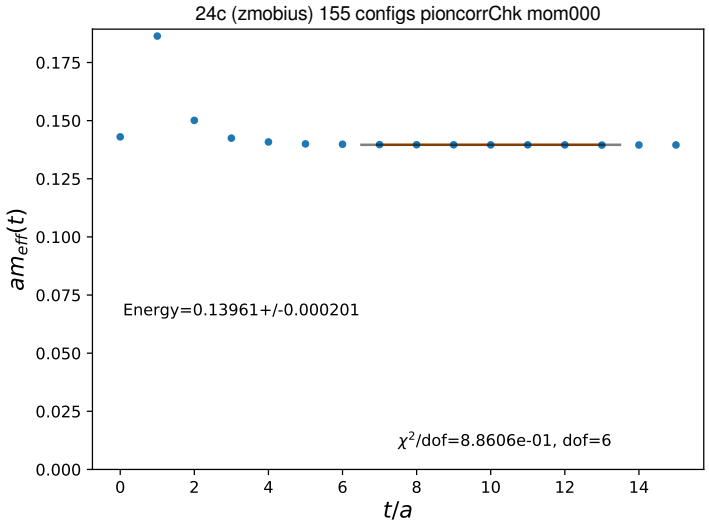
$$\cosh(E) = \cosh(m) + \sum_{i=1}^{3}(1 - cos(p_i)) \tag{10.2}$$

P	E (lattice units)	(E) eq. (10.2)	$\Delta E * 10^5$	stat.
m_π	0.13961(10)	-	-	-
$p1$	0.29612(15)	0.29502	-110	-
$p11$	0.39431(22)	0.39222	-209	-
$p111$	0.47223(45)	0.46894	-329	-

Table 10.1: A different dispersion relation

This appears to not work well for QCD, but maybe works well for the scalar field theory they were studying.

10.4 Effective Mass Plots



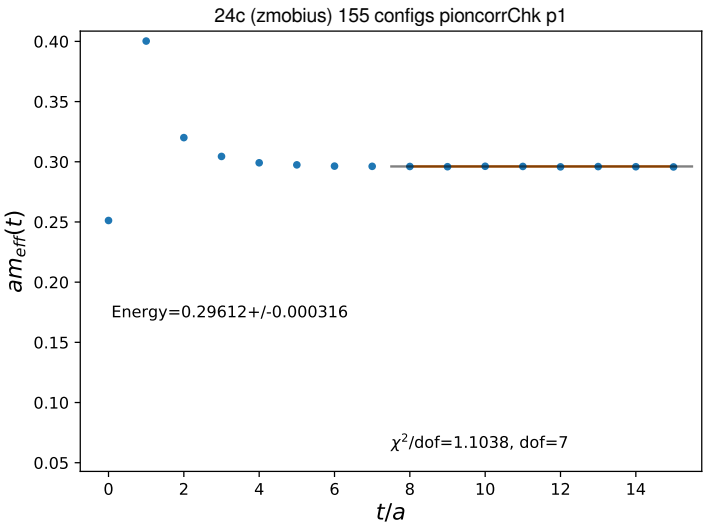
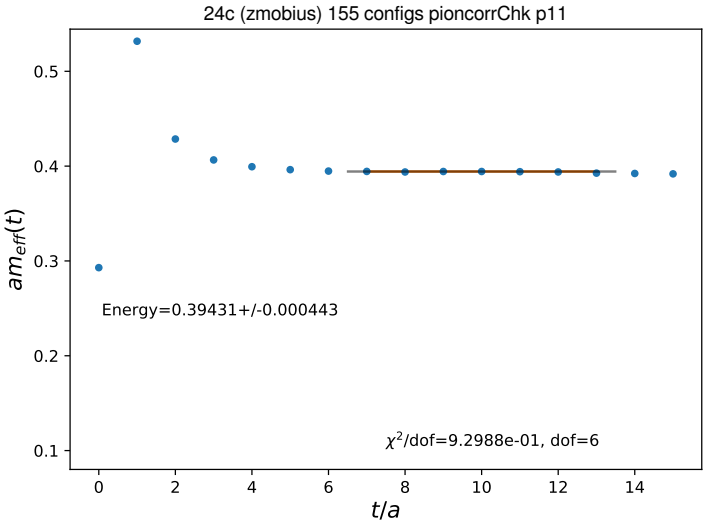
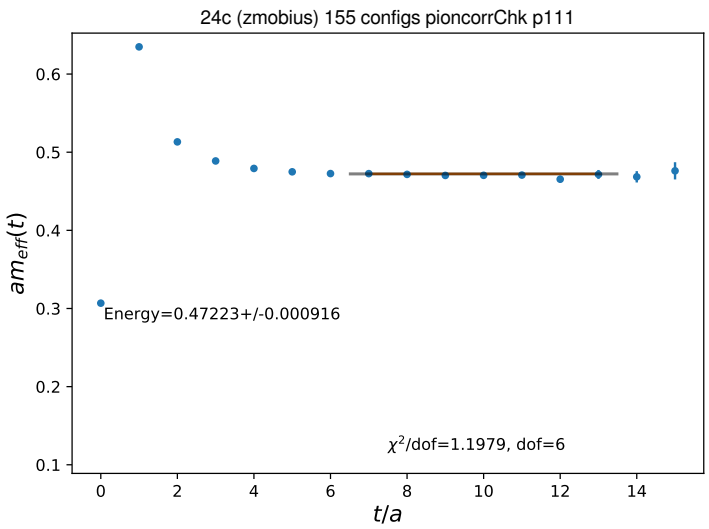


Figure 10.1: The fit range is slightly different because the original fit range gives a singular covariance matrix.





10.5 Discretization Error Corrections

We measure the amount of discretization error via differences between the continuum Einstein relativistic dispersion relation and fits to single pions. Once a significant discrepancy has been found, various options exist for correcting it back to the continuum relation. These measures are considered to be a stop-gap until one can afford to run on a finer lattice to take the continuum limit directly. In addition to the methods listed in this section, we also can correct the discretization error via section 8.10.

10.5.1 Additive Correction to $E_{\pi\pi}$

If the two particle channel has no resonances, we can suppose that, even if the discretization error varies a lot for different energies, the differences between the sum of single particle energies and the sum of the sum of continuum dispersion relation energies will correct the interacting energies (which lie in the vicinity of the free energies). We can get the mass used in the dispersion relation in various ways. For the periodic case, we simply fit to the stationary single pion correlation function.

Thus, a correction to the $p1$ ground state will be composed as, e.g.,

$$E_\pi(p1)+E_\pi(0)-\sqrt{m^2+(\frac{2\pi}{L})^2}-m=E_\pi(p1)-\sqrt{m^2+(\frac{2\pi}{L})^2}$$

This correction is implicitly applied in the method of section 8.10. We used to apply this to $I=2$, since the interaction energies were small. N.B., the corrections should be fully jackknifed in order to get a true jackknife estimate of the energies.

10.5.2 Lattice Dispersion Relation: $\hat{p}$

We can change the continuum dispersion relation in all equations to

$$E = \sqrt{m^2 + p^2} \rightarrow E = \sqrt{m^2 + \hat{p}^2}$$
$$\hat{p} \equiv 2sin\left(\frac{\pi}{L}\right)$$

In the continuum limit, $L \rightarrow \infty$ and we recover the continuum relation. Before the pion ratio method was seen to give decent results in $I = 0,1$, we applied this (ad-hoc) procedure to $I = 0,1$ to account for discretization errors which might occur (esp. in section 3.2).

10.6 Additional Hits - 24^3 Results

We look at the effective mass with errors for several operators on 10 (sloppy only) configurations of 24^3 (see section 3.2) data. This data comes from the spin contaminated set (with missing correlation functions) found under /cache/K2pipiPBC/qcddata/DWF/2+1f/24nt64/IWASAKI+DSDR/b1.633/ls24/M1.8/ms0.0850/ml0.00107/pipi at Thomas Jefferson National Laboratory (hereafter referred to as Jlab). There are folders labeled "1hitplus,2hitplus,3hitplus" which have data generated by incrementing the random number seed. One can also verify this study (possible future work) using spin uncontaminated data found under

/cache/K2pipiPBC/qcddata/DWF/2+1f/24nt64/IWASAKI+DSDR/b1.633/ls24/M1.8/ms0.0850/ml0.00107/new_pipi/job-0*1

(that is, the jobs end in 1 since we've incremented the random number seed by one). We use the log of the ratio of eigenvalues at adjacent times (see eq. (8.28)).

Table 10.2: $\langle\sigma|\sigma\rangle$ Hits Results

Time	1 Hit	2 hits	3 hits
0	0.394(63)	0.381(55)	0.419(46)
1	0.75(15)	0.750(128)	0.729(127)
2	0.627(275)	0.626(215)	0.618(197)
3	0.56(79)	0.191(531)	0.639(490)
4	NaN	-	-

$\langle\sigma|\sigma\rangle$ effective mass in lattice units, "–" indicates a data omission; NaN means a negative log argument.

Hits Results for $\langle \rho | \rho \rangle$

Time	1 Hit	2 hits	3 hits
0	0.043(93)	0.0784(907)	0.055(93)
1	1.05(6)	1.03(7)	1.01(6)
2	0.856(64)	0.886(59)	0.870(44)
3	0.825(105)	0.785(79)	0.789(74)
4	0.776(135)	0.82(10)	0.812(96)
5	0.705(150)	0.743(81)	0.702(77)
6	0.639(117)	0.640(123)	0.651(105)

Table 10.3: ρ Hits Results

As one can see, the ρ appears to display $\sqrt{N}$ scaling out to 2 hits, whereas the σ displays $\sqrt{N}$ scaling even out to 3 hits. This motivated our upcoming study of an additional hit for all configurations of the 24^3. We also looked at the $\pi\pi$ operator with one unit of back-to-back momentum, but it did not seem to improve with additional hits.

10.7 Phase Shifts

We plot phase shifts vs. χ_{pt} (chipt) predictions (see [13]) and vs. phenomonelogy (pheno) predictions (see [14]).

10.7.1 $\pi\pi$ $I = 2$

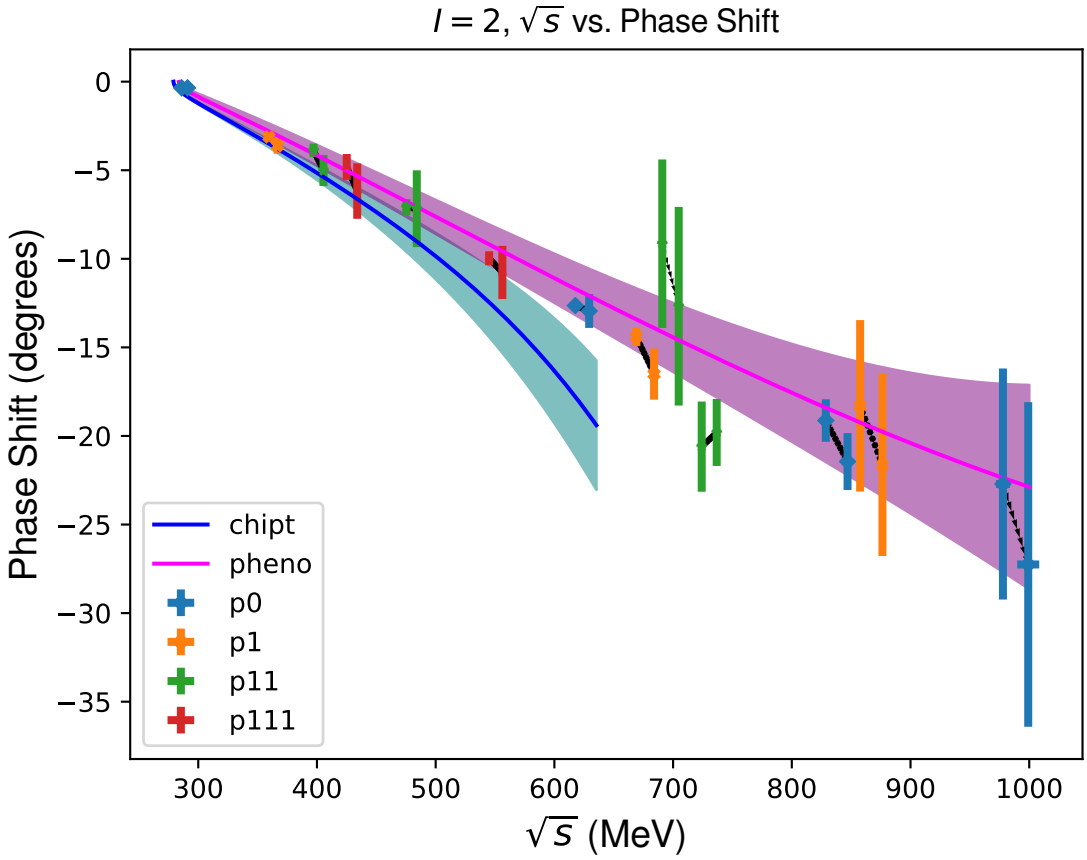


Figure 10.2: $I = 2$ phase shifts

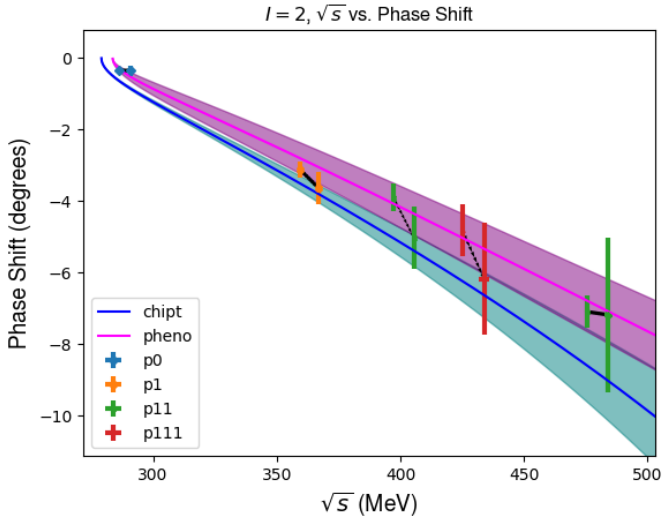


Figure 10.3: $I = 2$ phase shifts, zoomed

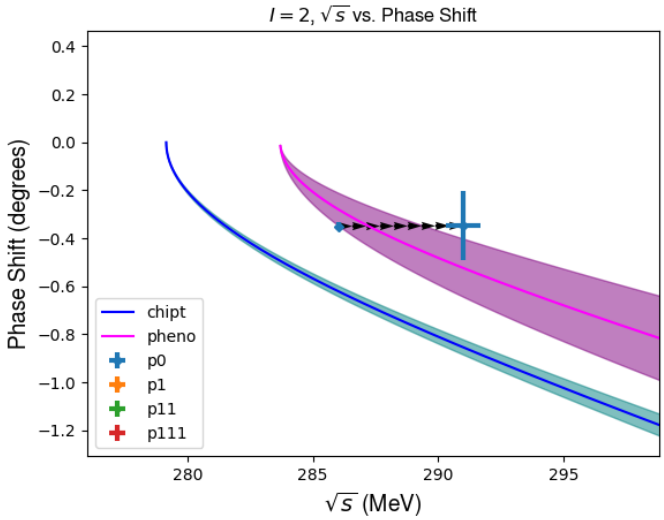


Figure 10.4: $I = 2$ phase shifts, zoomed (2)

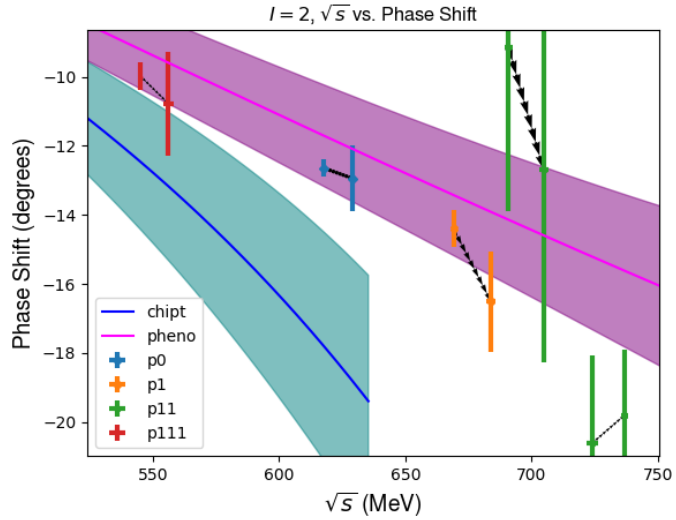


Figure 10.5: $I = 2$ phase shifts, zoomed (3)

For prior work, see [1][2].

Comments: The phase shifts for 24^3 are always to the left of the 32^3 due to the fact that the latter has a slightly larger pion mass. Arrows are drawn to indicate that the continuum limit is towards the 32^3. However, a continuum limit has not yet been performed. GEVP effective mass fits (and explicit numerical values) for figs. 10.2 to 10.5 are displayed in section A.3. Phase shifts and energies are pvalue weighted averages of $O(16)$ fit ranges which are pseudo-randomly selected. We fit to a constant, and increase tmin until all fit ranges give pair-wise consistent results (within $\sim 1.5\sigma$). Excited state contamination tends to make the energies larger,

which makes the phase shifts more negative. Roy equations, see curves labeled "pheno") analysis roughly involve interpolating between fits to low and high energy data. Thus, these phase shifts tend to be reliable at low and high ends of our scale. Chipt involves computing a finite number of higher dimensional operators via perturbation theory, so tends to be reliable at lower energies (which is why we cut off the chipt curve before it diverges).

Based on experimental data so far obtained, it is well-known that we don't have a resonance in this channel. The bulk energy of states we usually compute is the non-interacting energy. We thus see a correspondence between non-interacting and free energies. A resonance is an extra state inserted between two free (non-interacting) states, so we can see that it forces us to pass through a phase shift of π. Without a resonance, we can trust extrapolation from chipt and interpolation from the Roy equations. Since this channel is thus "clean" (of resonances), we can use this to perfect methodology and understandings we can apply to $I = 0, 1$. Additionally, $I = 2$ states are clean since they not very noisy statistically.

While our stable plateaus might encourage us to believe we are free from excited state contamination, we might easily be in the regime of section 8.2.2. Thus, hypothetically, the contaminating excited state comes from a radially excited pion (with the same corresponding $\pi\pi$ momenta). Future work might test this hypothesis via a more sophisticated GEVP analysis or by adding additional operators.

Other sources of error include discretization errors. However, since we always use the pion ratio method (section 8.10), we automatically remove some of this error.

We are also well beyond the four-π threshold (of $O(500)MeV$), but we suppose that the lack of phase space suppresses these amplitudes (which, based on lattice studies of these amplitudes as well as experimental data, is not an unreasonable assumption).

10.7.2 $\pi\pi$ $I = 1$

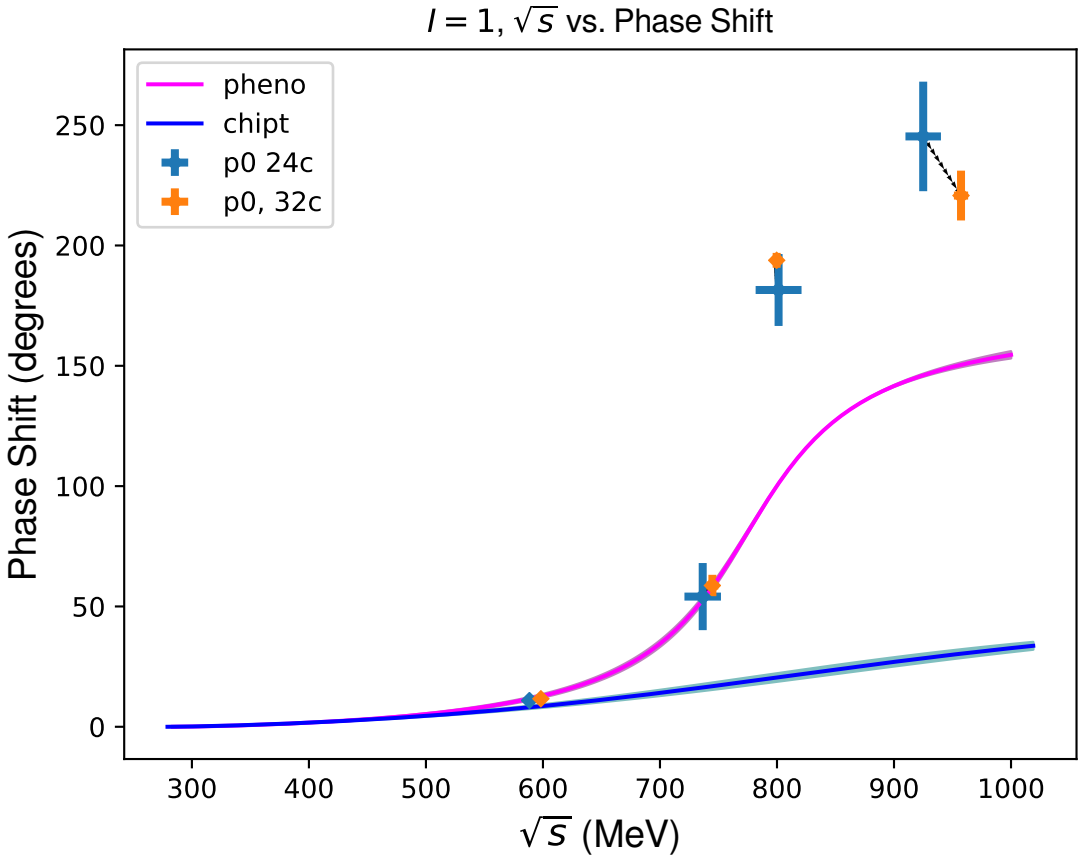


Figure 10.6: $I = 1$ phase shifts

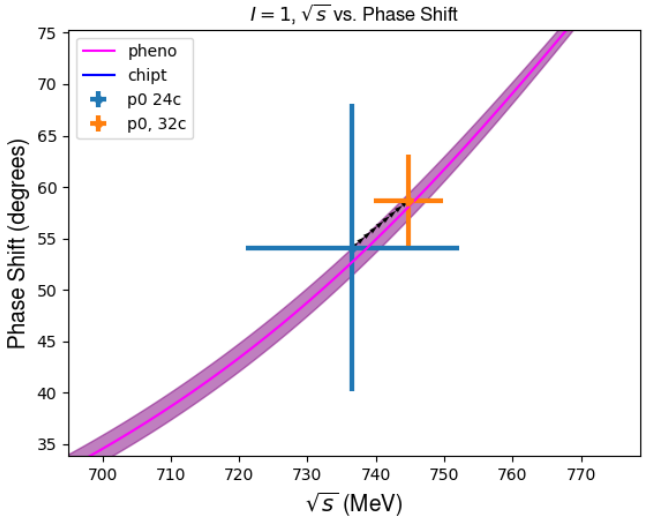


Figure 10.7: $I = 1$ phase shifts, zoomed

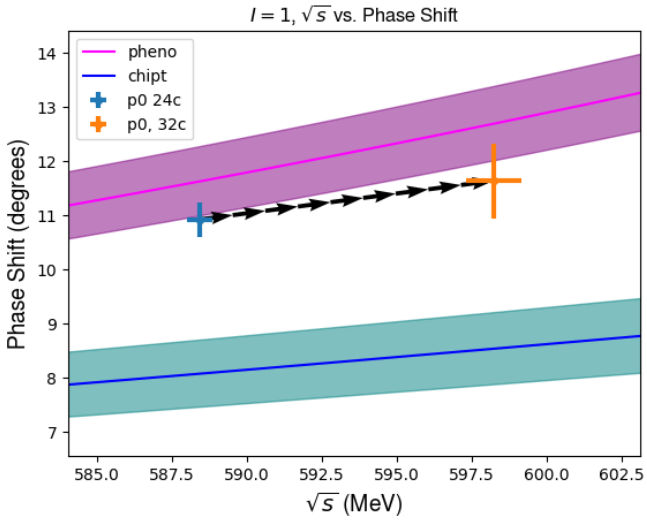


Figure 10.8: $I = 1$ phase shifts, zoomed (2)

For prior work, see [3][4].

Commentary: We fit to $O(1)$ fit range (given that fitting as in section 9.7 makes finding a fit range which gives a good χ^2 much more difficult than the constant fit case) for the 24^3, but use a systematic error ansatz (see section 9.7). The 32^3 is fit to a constant, although prior results (not included; fit to earlier times) did not find much difference from fits to a constant. We use the pion ratio method (although not using it seems (meaning not extensively tested) not to make a big difference). Clearly, we have now a resonance in our channel (which we suppose corresponds to the ρ particle at around 770 MeV). While the first two points have good agreement (with Roy equation analysis), the last two have a very large discrepancy. We know from section 8.1 the higher state excited state error will decay more slowly the higher the energy we are trying to resolve. This is consistent with the highest two states have large amounts of excited state contamination. In fact, if we have more than one resonance state (which is below our highest two $\pi\pi$ states), this might account for such a large discrepancy. An independent analysis of this data was performed by T. Izubuchi (priv. comm.), who found good agreement with fig. 10.6 (though he did not pvalue average as we did here, so one of the errors on the phase shift was $O(5)$ degrees different).

It is also important to note that although Lüscher's formula is ambiguous by π radians for the phase shift. However, by tracking the free vs. interacting energy correspondence, we can find whether a state is the result of smoothly passing 180 degrees, or if it actually is close to its corresponding non-interacting energy (in which case we might give it a 0 degree phase shift). Since we have a resonance, we know the higher states have very large phase shifts (although shifting down by 180 degrees does not bring these points into alignment with the Roy equations).

Another recent $\pi\pi$ $I = 1$ study (of A. Meyer; priv. comm.) has found that having a second $\overline{\psi}\gamma_\mu\psi$ operator can fix errors of this type (including very large phase shift systematics). However, given that we use the more expensive A2A method (than distillation; see section 2.0.1), it is not feasible to conduct such a study at this time.

The chipt curve does not have operators corresponding to the ρ resonance, so it is going to be very unreliable around the resonance energy. We include it in fig. 10.6 mainly to illustrate this point.

10.7.3 $\pi\pi$ $I = 0$

Prior work on $I = 0$ $\pi\pi$ (non-physical quark mass) can be found in [5].

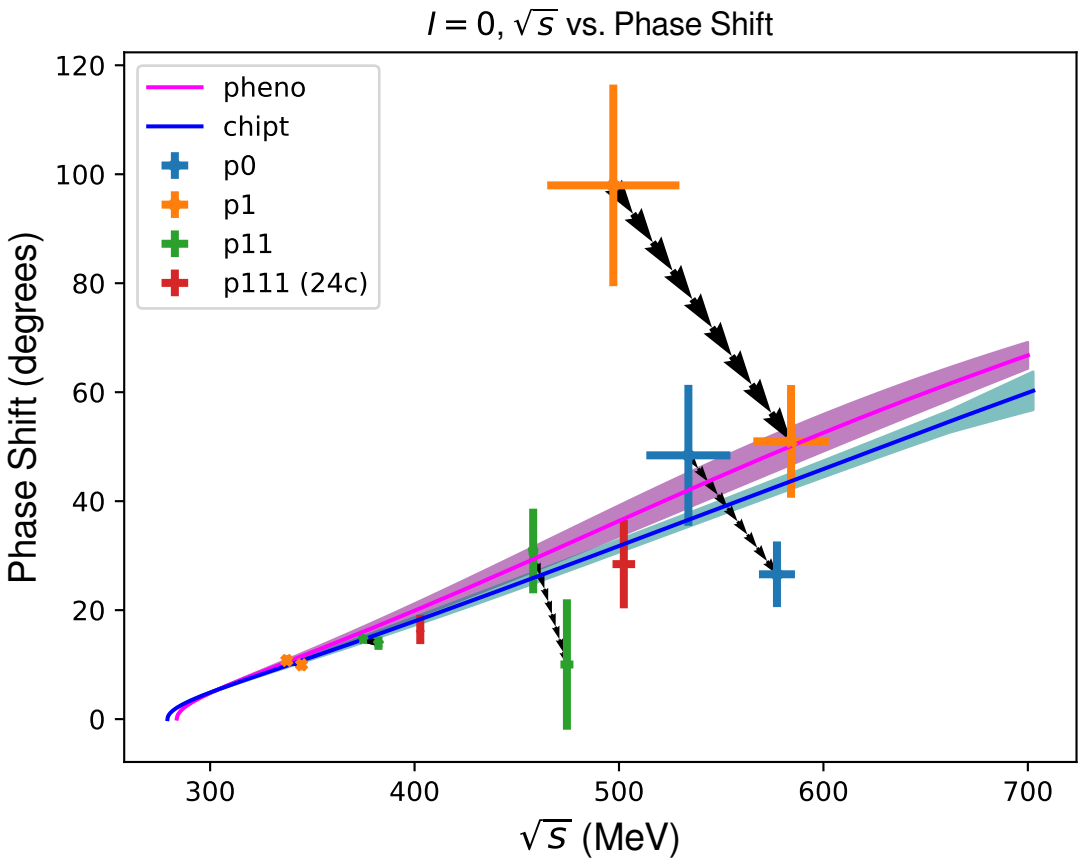


Figure 10.9: $I = 0$ phase shifts; p111 points are of the 24^3 due to the large defect (likely ATW) found in 32^3

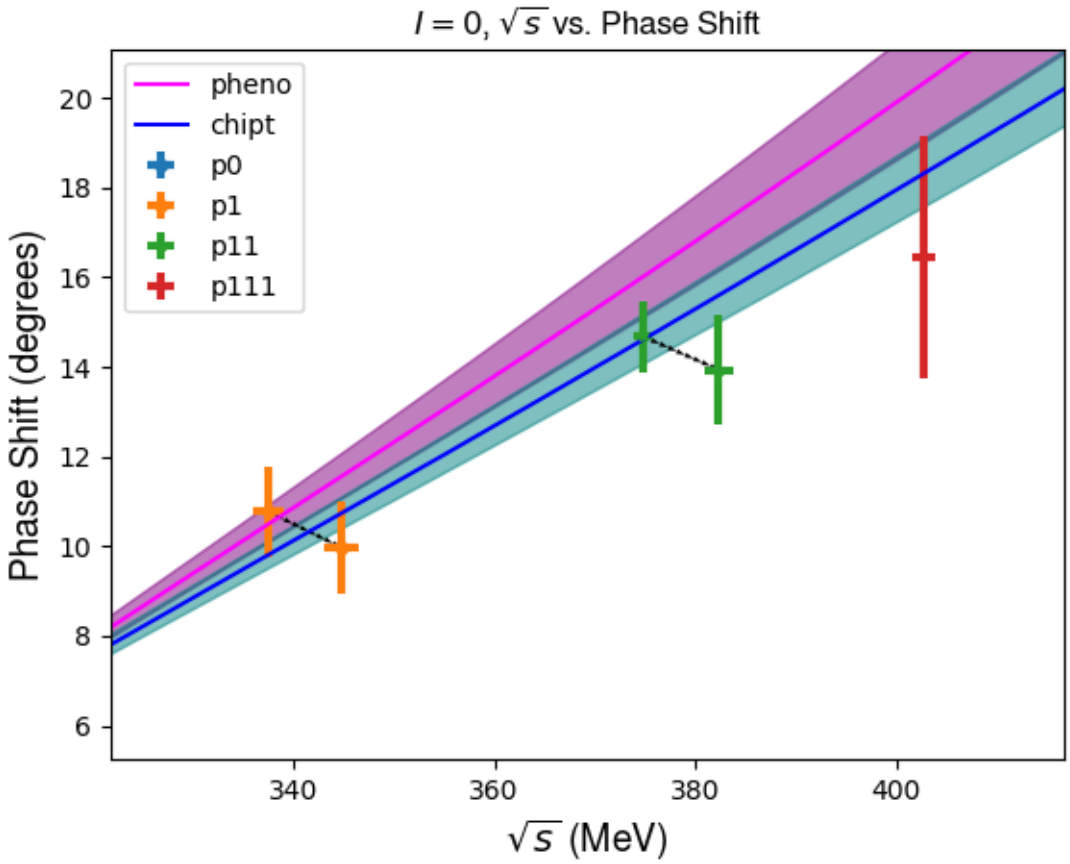


Figure 10.10: $I = 0$ phase shifts, zoomed

It is unclear why there is such a large jump in p1 from 24^3 to 32^3.

10.8 Coda: Around the World Subtraction

We close out chapter 10 with a brief discussion of an underestimated[1] systematic error: around the world terms. We first show, in graphical form, the difference between using eq. (8.22) and using the technique of section 8.4.2 used in fig. 10.2 followed by use of the method of section 9.7 with matrix subtraction to estimate the around the world terms.

[1] in the sense that we don't know what systematic error is associated with interacting around the world terms

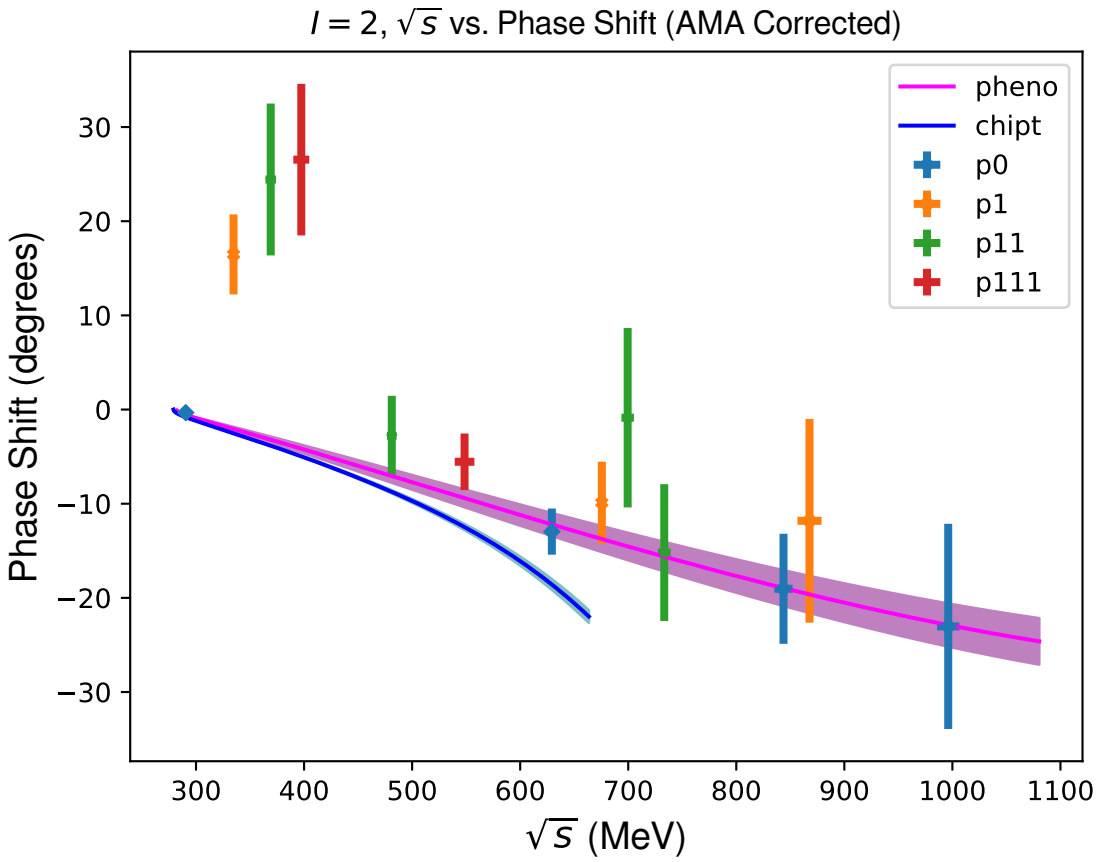


Figure 10.11: 32^3, $I = 2$, 65,8 (sloppy,exact) configs; with matrix subtraction

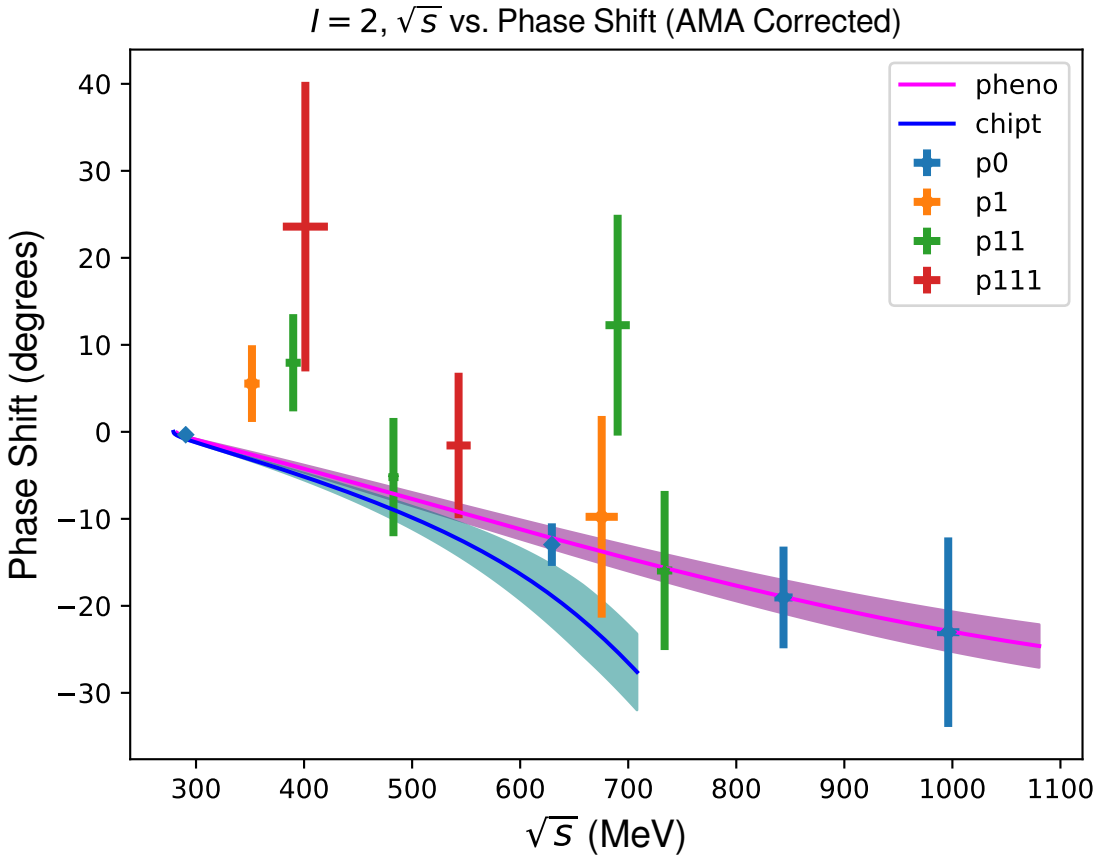


Figure 10.12: 32^3, $I = 2$, 65,8 (sloppy,exact) configs; with matrix subtraction and fit ansatz from section 9.7

The results are clear: for such a small physical time extent we need to do more than remove the two leading order non-interacting terms

10.8.1 A Difficult Counter-Example

Now we look at a difficult counter-example to the notion that vacuum saturation subtraction removes the majority of the ATW contamination.

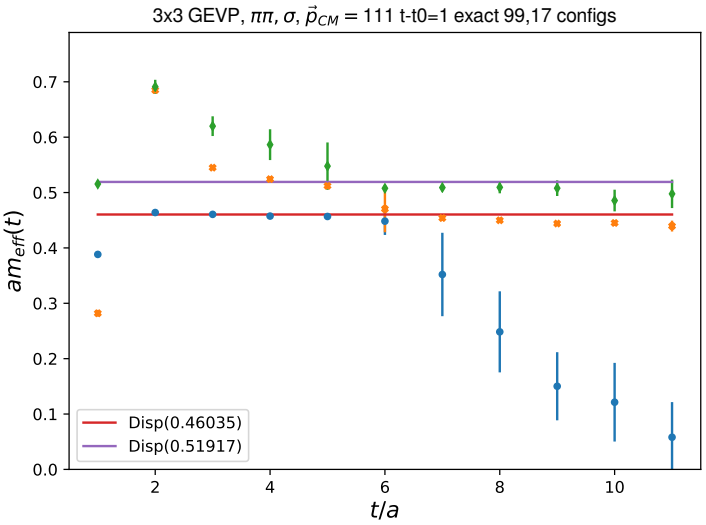


Figure 10.13: $I = 0$ p111, 32^3

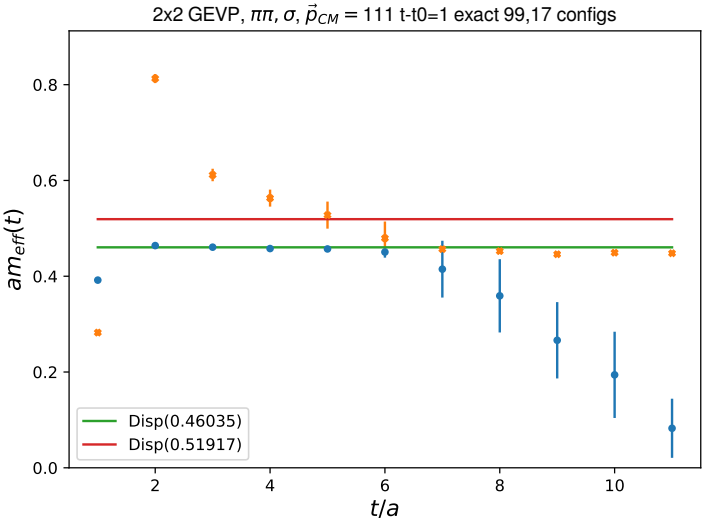


Figure 10.14: $I = 0$ p111, 32^3 without the top operator (pi(001)pi(011))

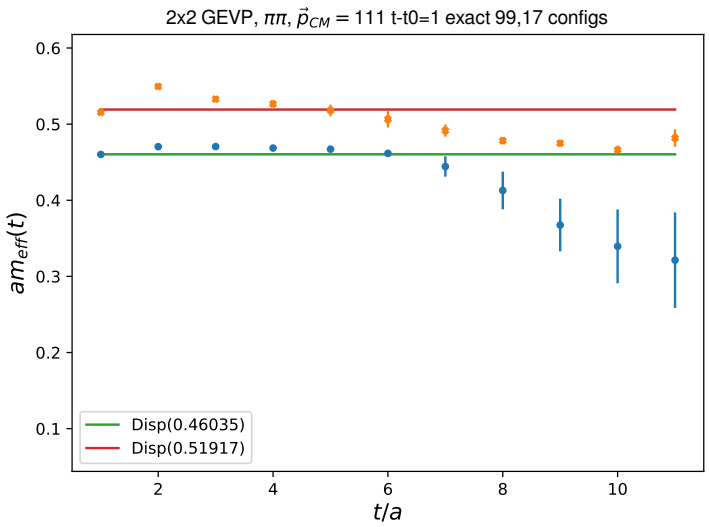


Figure 10.15: $I = 0$ p111, 32^3 without the σ operator

These results are consistent with a large ATW term. The σ operator usually has large overlaps with all states, so it is not surprising that around $t = 6$ we see a clear decay of one of the excited states to the (early time) ground state which is around the lowest dispersive line. In fig. 10.15, it is clear that there is much slower decay, which is consistent with a mostly diagonal GEVP. These plots follow the general methodology of the others, so use the methods of sections 8.4.2 and 8.10, although matrix subtraction (section 8.4.1) and no pion ratio does not help reduce the amount of ATW contamination. We know the smaller physical time extent of the 32^3 ensemble generates large ATW terms. It is an open question, however, as to why the terms are so much more likely to be interacting in the $p111$ case.

Chapter 11

$K \to \pi\pi$

We write down the process $K^0 \to \pi\pi$ in terms of standard (non-Gparity) operator expansion involving operators Q_i. The final state pions are projected onto $I = 0, 1, 2; I_3 = 0$. These contractions were derived without looking at [7]. They duplicate (in a briefer form) much of the content there. Additionally, preliminary analysis code for $K \to \sigma, K \to \pi\pi$ can be found in the module "kaonanalysis" in [15].

11.1 Diagram types

In general, for each term we have a three point function given in terms of times $t_{sep}, t_{dis}, t_{op}, t_{src}$. t_{sep} is the usual distance between pions at the sink (a computational technique to avoid overlap with the vacuum). We display the diagrams with non-zero t_{sep}. t_{dis} is the distance between the four-quark operators Q_i and the inner (sink) pion (the one closest to the operator in time). t_{op} is the time slice location of the four-quark operator. t_{src} is the time slice location of the K. We draw the diagrams with time flowing left to right, and arrows starting on $\overline{q}$ and ending on the q. To map these onto the contractions below, take the outer pion to also be the least inner in the bracket: $\langle 0|\pi_{\text{outer}}\pi_{\text{inner}}...|0\rangle$. We display the strange quark contraction as a double line. We assume the u, d quarks are degenerate (so we have no single pion bubbles).

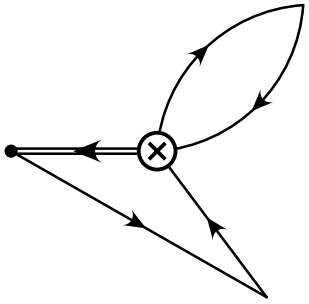
We label the contractions with letters. Different topologies of the same contraction are labelled with subscripts. The strange quark propagator is denoted S^{-1} instead of the usual M^{-1} for the u, d quarks. Primed vs. unprimed diagrams are defined in sections 11.1.1 to 11.1.4.

In the penguin diagrams, $V-A$ can be replaced with $V+A$ at the sink (because the sink only interacts with a gluon, photon, or Z at the operator). We know that it's the sink and not the source that gets the $V+A$ since only $V-A$ can change flavor (and we are changing strangeness in the $K \to \pi\pi$ process). We thus display $\pm$ in the internal operator spin structure of the sink. These diagrams should be understood to have the appropriate

sign depending on the context in which they are read.

Finally, some notation needs to be defined. The 4-quark operator has two fermion loops connected to it. Each of these loops could be a spin trace, or we could have an overall spin trace of both loops. This is the difference between the primed and unprimed diagram letters (e.g. f_1 vs. f_1'). These spin traces could also be color traces, in which case $Tr(...)$ refers to both and we simply refer to the figure by the letters defined in sections 11.1.1 to 11.1.4. The other possibility is that the color trace takes the opposite (primed vs. unprimed) structure vs. the structure of the spin trace. Let x, x' be a $K \to \pi\pi$ diagrams. We denote the situation where the opposite trace structure should be used for the color trace by, respectively, $Tr_c(x), Tr_c(x')$.

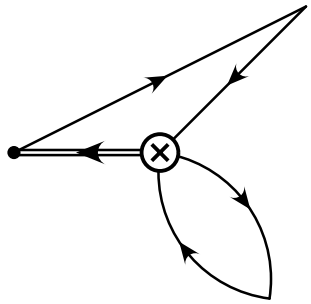
11.1.1 Type 1 Diagrams



$$\equiv f_1, f_1'$$

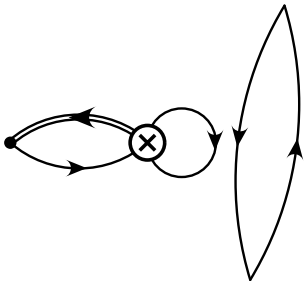
$$f_1 = Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$

$$f_1' = Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}\gamma_5 M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$



$$\equiv f_2, f_2'$$

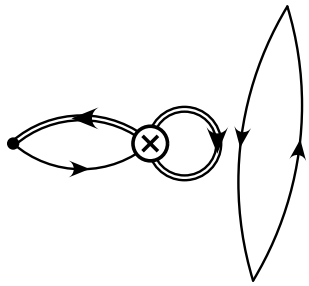
11.1.2 Type 2 Diagrams



$$\equiv g, g'$$

$$g = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5\right)$$

$$g' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5\right)$$

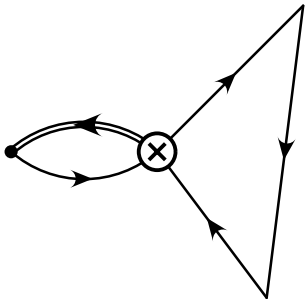


$$\equiv G, G'$$

$$g = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)\right)\cdot Tr\left(M^{-1}\gamma^\mu(1\pm\gamma_5)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5\right)$$

$$g' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)M^{-1}\gamma^\mu(1\pm\gamma_5)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5\right)$$

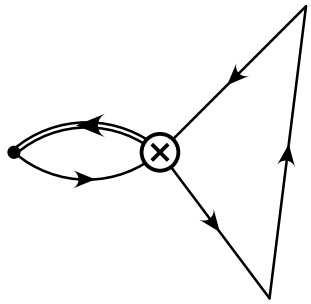
11.1.3 Type 3 Diagrams



$$\equiv h_1, h_1'$$

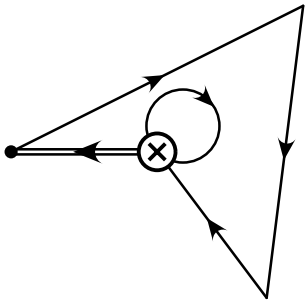
$$h_1 = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)\right)\cdot Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma^\mu(1\pm\gamma_5)\right)$$

$$h_1' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma^\mu(1\pm\gamma_5)\right)$$



$$\equiv h_2, h_2'$$

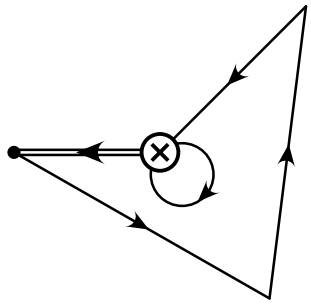
11.1.4 Type 4 Diagrams



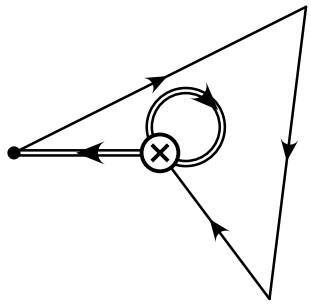
$$\equiv i_1, i_1'$$

$$i_1 = Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)\right)\cdot Tr\left(M^{-1}\gamma^\mu(1\pm\gamma_5)\right)$$

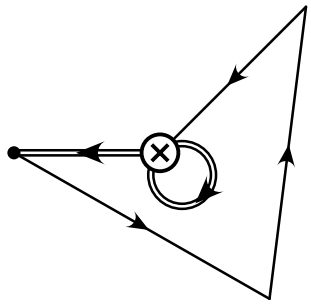
$$i_1' = Tr\left(M^{-1}\gamma_5 M^{-1}\gamma_5 M^{-1}\gamma_5 S^{-1}\gamma_\mu(1-\gamma_5)M^{-1}\gamma^\mu(1\pm\gamma_5)\right)$$



$$\equiv i_2, i_2'$$



$\equiv I_1, I_1'$



$\equiv I_2, I_2'$

11.2 Analysis

If we ignore the differences between the $1,2$ diagrams[1], and have perfect $SU(3)$ symmetry[2],we get 32 diagrams

32 =(4: diagram types)(2: left vs. right handed sink)(2: primed vs. unprimed)(2: x vs. $Tr_c(x)$).

Also, N.B. if x has $+1$ from Grassmann number commutation, x' has -1 (and vice-versa).

All correlation functions follow the general structure:

$$\langle \pi\pi | Q_i | K \rangle$$

11.3 Current-Current Operators

11.3.1 Q_1: Color Diagonal

$$\begin{aligned}\left\langle \pi\pi | (\overline{u}_\beta u_\beta)_{V-A} (\overline{s}_\alpha d_\alpha)_{V-A} | K \right\rangle &= \left\langle \pi\pi | (\overline{u}_\beta u_\beta)_{V-A} (\overline{s}_\alpha d_\alpha)_{V-A} \overline{d}\gamma_5 s | 0 \right\rangle \\ &= \left\langle \pi\pi | (\overline{u}_\beta u_\beta)_{V-A} (\overline{s}_\alpha d_\alpha)_{V-A} \overline{d}\gamma_5 s | 0 \right\rangle\end{aligned}$$

We drop the γ and $V \pm A$ factors in what follows for the sake of brevity.

$$= \left\langle \pi\pi | (\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s | 0 \right\rangle$$

[1]These are the same up to swapping of the sink pions; we average over these two topologies. For completeness, we show the intermediate sum with both topologies, and when we refer to this average when we drop the subscript in the final step.

[2]We list capital letters for the case of SU(2) symmetry, and denote the strange quark propagator by a double line.

11.3.1.1 $I = 0$

Per chapter 4, we have the expansion for $\pi\pi_{I=0}$:

$$|\pi\pi\rangle_{I=0} = \frac{1}{\sqrt{3}}\left(u\bar{d}d\bar{u} + \frac{1}{2}\left(u\bar{u}u\bar{u} - u\bar{u}d\bar{d} - d\bar{d}u\bar{u} + d\bar{d}d\bar{d}\right) + d\bar{u}u\bar{d}\right)$$

$$\Rightarrow \left\langle \pi\pi_{I=0}|(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s|0\right\rangle = \left\langle 0|\frac{1}{\sqrt{3}}\left(\bar{d}u\bar{u}d + \frac{1}{2}\left(\bar{u}u\bar{u}u - \bar{u}u\bar{d}d - \bar{d}d\bar{u}u + \bar{d}d\bar{d}d\right) + \bar{u}d\bar{d}u\right)(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s|0\right\rangle$$

This gives us six terms. Let's work them out one by one.

1.

$$\begin{aligned}\frac{1}{\sqrt{3}}\left\langle 0|\bar{d}u\bar{u}d(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s|0\right\rangle &= \frac{1}{\sqrt{3}}\left\langle 0|\bar{d}u\bar{u}d(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}_\delta s_\delta|0\right\rangle \\ &= \frac{1}{\sqrt{3}}\left(i_2 + h_1 - g - f_1'\right)\end{aligned}$$

2.

$$\frac{1}{2\sqrt{3}}\left\langle\left(\bar{u}u\bar{u}u(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s\right)\right\rangle = \frac{1}{2\sqrt{3}}(h_1 + h_2 - g)$$

3.

$$-\frac{1}{2\sqrt{3}}\left\langle\left(\bar{u}u\bar{d}d(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s\right)\right\rangle = -\frac{1}{2\sqrt{3}}f_1$$

4.

$$-\frac{1}{2\sqrt{3}}\left\langle\left(\bar{d}d\bar{u}u(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s\right)\right\rangle = -\frac{1}{2\sqrt{3}}f_2$$

5.

$$\frac{1}{2\sqrt{3}}\left\langle\left(\bar{d}d\bar{d}d(\bar{u}_\beta u_\beta)(\bar{s}_\alpha d_\alpha)\bar{d}s\right)\right\rangle = \frac{1}{2\sqrt{3}}(i_1 + i_2 - g)$$

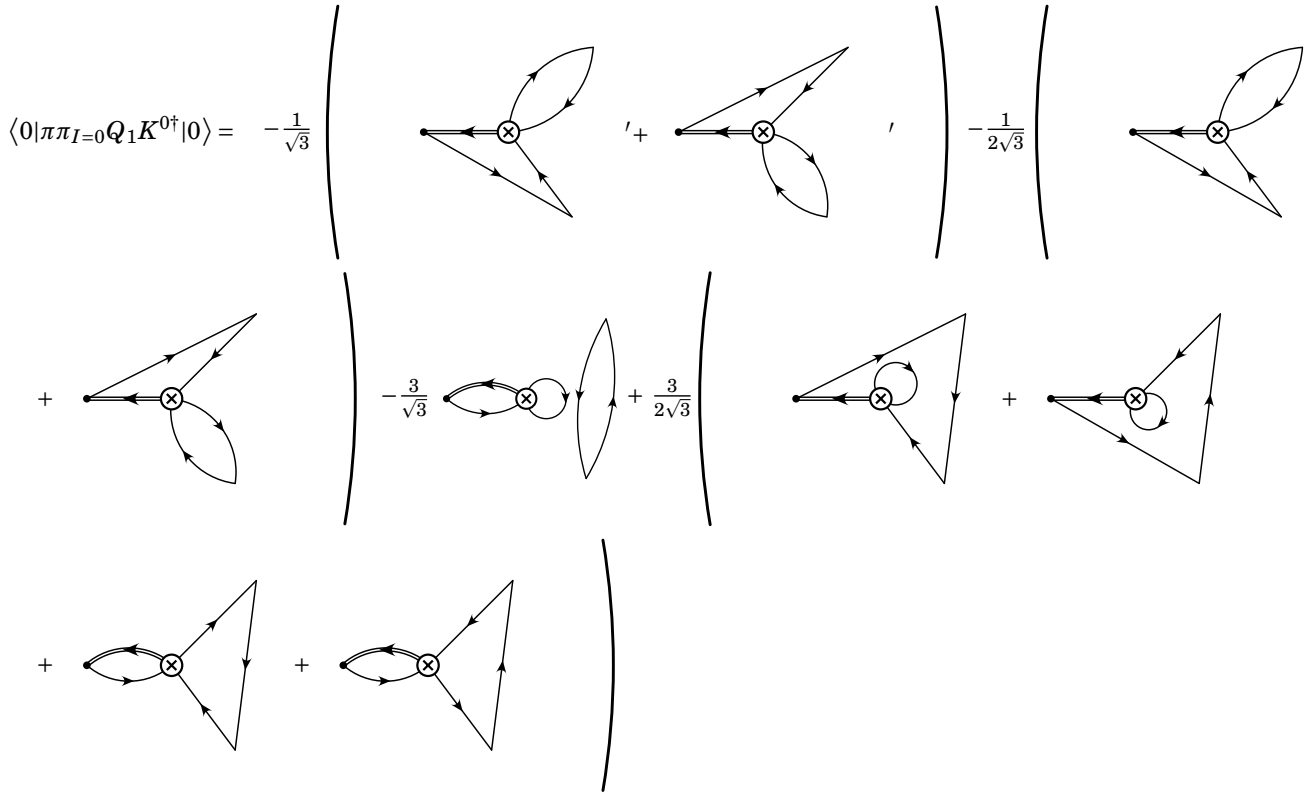
6.

$$\frac{1}{\sqrt{3}}\left\langle 0|\overline{u}d\overline{d}u(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}_\alpha s_\alpha|0\right\rangle = \frac{1}{\sqrt{3}}\left(h_2+i_1-g-f_2'\right)$$

Finally, the total is

$$\begin{aligned}\left\langle 0|\pi\pi_{I=0}Q_1K^{0\dagger}|0\right\rangle &= \frac{1}{\sqrt{3}}\left(\left(i_2-g+h_1-f_1'\right)+\left(i_1-g+h_2-f_2'\right)\right)-\frac{1}{2\sqrt{3}}\left((g-i_1-i_2)+f_1+f_2+(g-h_1-h_2)\right)\\ &= -\frac{1}{\sqrt{3}}\left(f_1'+f_2'\right)-\frac{1}{2\sqrt{3}}(f_1+f_2)-\frac{3}{\sqrt{3}}g+\frac{3}{2\sqrt{3}}(i_1+i_2+h_1+h_2)\\ &= -\frac{2}{\sqrt{3}}f'-\frac{1}{\sqrt{3}}f-\frac{3}{\sqrt{3}}g+\frac{3}{\sqrt{3}}(i+h)\\ &= -\frac{1}{\sqrt{3}}\left(f+2f'+3g-3(i+h)\right)\end{aligned}$$

$\langle 0|\pi\pi_{I=0}Q_1K^{0\dagger}|0\rangle =$



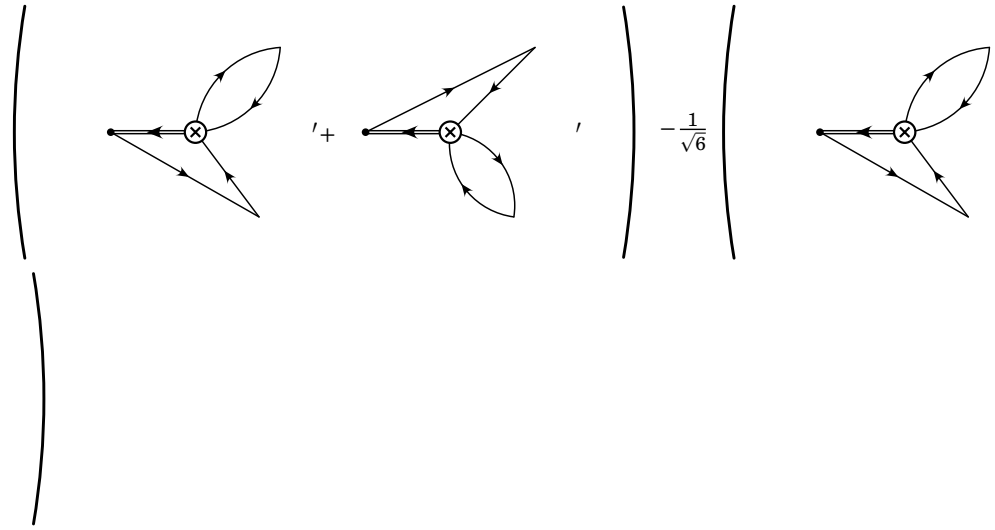
11.3.1.2 $I=2$

Per chapter 4, we have the expansion for $\pi\pi_{I=2}$:

$$|\pi\pi\rangle_{I=2} = \frac{1}{\sqrt{6}}\left(-\overline{u}d\overline{d}u + \overline{u}u\overline{u}u - \overline{u}u\overline{d}d - \overline{d}d\overline{u}u + \overline{d}d\overline{d}d - \overline{d}u\overline{u}d\right)$$

$$\begin{aligned}\left\langle 0|\pi\pi_{I=2}Q_1K^{0\dagger}|0\right\rangle &= \frac{1}{\sqrt{6}}\left(-\left(i_2 - g + h_1 - f_1'\right) - \left(i_1 - g + h_2 - f_2'\right) - f_1 - f_2 - 2g + h_1 + h_2 + i_1 + i_2\right)\\ &= \frac{1}{\sqrt{6}}\left(f_1' + f_2' - f_1 - f_2\right)\\ &= \frac{2}{\sqrt{6}}\left(f' - f\right)\end{aligned}$$

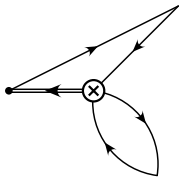
$\left\langle 0|\pi\pi_{I=2}Q_1K^{0\dagger}|0\right\rangle =$ $\frac{1}{\sqrt{6}}$



$'+$

$'$ $-\frac{1}{\sqrt{6}}$

$+$



11.3.1.3 $I = 1$

We could write down these contractions, if the Kaon didn't have spin 0. Bose symmetry thus rules out all $I = 1$ diagrams.

11.3.2 Q_2: Mixed Color

Since the only flavor changing weak current is a $V - A$ current, these are the only current-current operators. We get this operator by exchanging color indices on the quarks in the 4-quark operator. Thus, any terms that are operated on by Tr_c are now non Tr_c diagrams and vice-versa. We write, trivially:

$$\left\langle 0|\pi\pi_{I=0}Q_2K^{0\dagger}|0\right\rangle = Tr_c(Q_{1,I=0})$$

$$\left\langle 0|\pi\pi_{I=2}Q_2K^{0\dagger}|0\right\rangle = Tr_c(Q_{1,I=2})$$

The color mixed operators are always operated on by Tr_c relative to their color-diagonal counterparts (hence we will omit writing them down in sections 11.4 and 11.5).

11.4 QCD Penguin Operators

The only other distinct dimension= 4 quark operators[3] can be found by going beyond tree level. These are the penguin diagrams. Depending on whether the penguin has a gluon or a photon/Z, we get two distinct classes of diagrams. We examine the ones involving gluons in this section.

11.4.1 Q_3: $V-A \to V-A$, Color Diagonal

$$\begin{aligned}\left\langle 0|\pi\pi Q_3 K^{0\dagger}|0\right\rangle &= \sum_{q=u,d,s}\left\langle 0|\pi\pi\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle \\ &= \sum_{q=d,s}\left\langle 0|\pi\pi\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle + Q_1\end{aligned}$$

11.4.1.1 $I=0$

Let's calculate each of the two terms separately.

We get all of $Q_1, I=0$ because $\pi\pi_{I=0}$ is symmetric with respect to d,u interchange (any contraction involving $u,\overline{u}$ from the operator and a pion now involves $d,\overline{d}$ from the operator, and a different term in $\pi\pi_{I=0}$ with $u \leftrightarrow d$).

We also get additional terms from contractions of the operator with the K/itself. These contractions weren't available when the operator had a $\overline{u}u$ (as in sections 11.3.1 and 11.3.2). The isospin coefficients don't change within a given diagram type for each of these additional possibilities (as these new contractions involve only the operator and the K. Thus, for each one, once we do this new contraction the $\pi\pi$ sink sees the exact same quark content so gives the exact same coefficient), so we can read them off from section 11.3.1 (once we know the diagram type, up to the Grassmann sign). One possibility is the q from the $\overline{q}q$ contracting with K, and the other possiblity is the $\overline{q}q$ contracting with the $\overline{s}d$ from the operator.

[3] To see these exhaust the available operators, consider that one of our two fermion lines (meaning a $\overline{q}q$ in the operator's quark content) is determined by needing to change strangeness. Then, either the other fermion line makes contact with a W, or it doesn't. Suppose it does. If this W originated on the distinguished fermion line, we fix the flavors of the other line, and get a current-current contribution. If it originated on the undistinguished line, then it doesn't do anything to distinguish this line's flavors. Also, there can be no initial to final state flavor change. Now, if this line doesn't make contact with a W, it must contact the operator through a Z, g, γ. We thus arrive at the penguin operators.

$$\left\langle 0|(\pi\pi)_{I=0}\overline{d}_\beta d_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle = Q_{1,I=0} + \frac{3}{\sqrt{3}}\left(g' - \frac{1}{2}\left(i'_1 + i'_2 + h'_1 + h'_2\right)\right)$$
$$\left\langle 0|(\pi\pi)_{I=0}\overline{s}_\beta s_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle = -\frac{3}{\sqrt{3}}\left(g - \frac{1}{2}(i_1+i_2)\right) + \frac{3}{\sqrt{3}}\left(g' - \frac{1}{2}\left(i'_1+i'_2\right)\right)$$

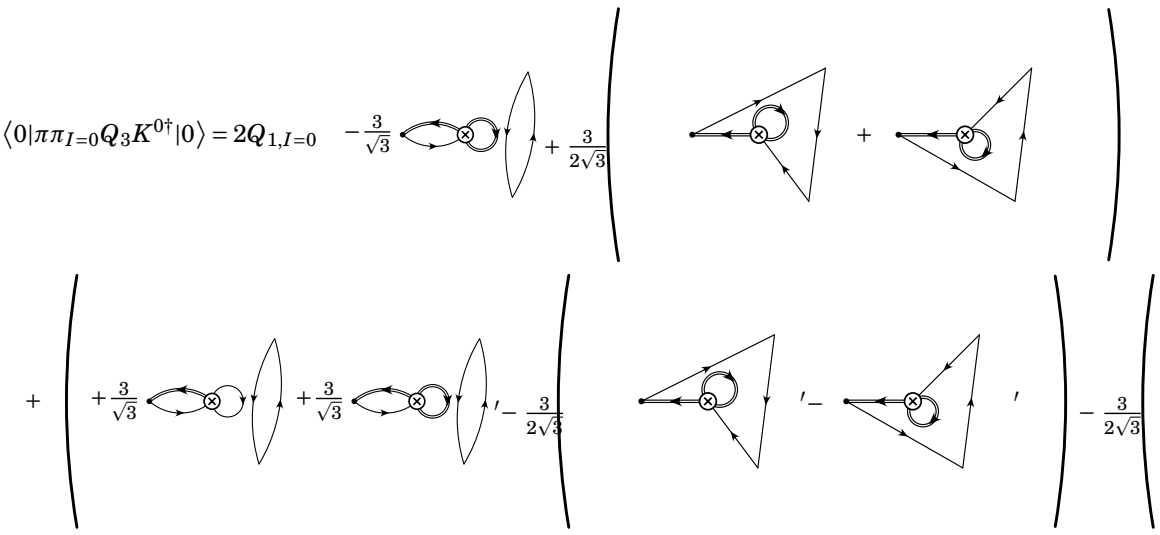
The s term must have a loop on the operator. Therefore, all terms from $Q_{1,I=0}$ that don't have this loop are excluded. This gives us finally,

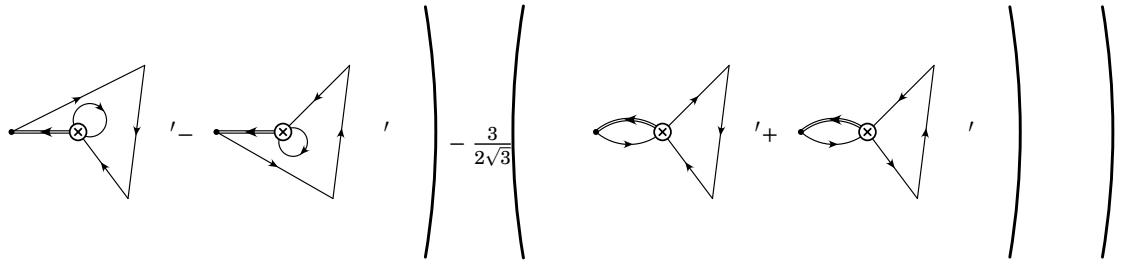
$$\left\langle 0|\pi\pi_{I=0}Q_3K^{0\dagger}|0\right\rangle = 2Q_{1,I=0} - \frac{3}{\sqrt{3}}\left(g - \frac{1}{2}(i_1+i_2)\right) + \frac{6}{\sqrt{3}}\left(g' - \frac{1}{2}\left(i'_1+i'_2\right) - \frac{1}{4}\left(h'_1+h'_2\right)\right)$$
$$\begin{aligned}\left\langle 0|\pi\pi_{I=0}Q_3K^{0\dagger}|0\right\rangle &= -\frac{2}{\sqrt{3}}\left(f'_1+f'_2\right) - \frac{1}{\sqrt{3}}(f_1+f_2) - \frac{6}{\sqrt{3}}g + \frac{3}{\sqrt{3}}(i_1+i_2+h_1+h_2) - \frac{3}{\sqrt{3}}\left(g - \frac{1}{2}(i_1+i_2)\right)\\ &\quad + \frac{6}{\sqrt{3}}\left(g' - \frac{1}{2}\left(i'_1+i'_2\right) - \frac{1}{4}\left(h'_1+h'_2\right)\right)\\ &= \frac{6}{\sqrt{3}}\left(-\frac{1}{3}(f'_1+f'_2) + g' - \frac{1}{2}\left(i'_1+i'_2\right) - \frac{1}{4}\left(h'_1+h'_2\right)\right) - \frac{1}{\sqrt{3}}(f_1+f_2) - \frac{9}{\sqrt{3}}\left(g - \frac{1}{2}(i_1+i_2)\right) + \frac{3}{\sqrt{3}}(h_1+h_2)\\ &= \frac{6}{\sqrt{3}}\left(-\frac{2}{3}f' + g' - i' - \frac{1}{2}h'\right) - \frac{2}{\sqrt{3}}f - \frac{9}{\sqrt{3}}(g-i) + \frac{6}{\sqrt{3}}h\end{aligned}$$

Distinguishing the strange quark propagator, we find

$$= \frac{3}{\sqrt{3}}\left(-\frac{4}{3}f' + g' - i' - h'\right) - \frac{2}{\sqrt{3}}f - \frac{6}{\sqrt{3}}(g-i) - \frac{3}{\sqrt{3}}\left(G - G' + I' - I\right) + \frac{6}{\sqrt{3}}h$$

$\left\langle 0|\pi\pi_{I=0}Q_3K^{0\dagger}|0\right\rangle = 2Q_{1,I=0}$





11.4.1.2 $I = 2$

Similar reasoning gives us:

$$\left\langle 0|(\pi\pi)_{I=2} Q_3 K^{0\dagger}|0 \right\rangle = 2 Q_{1,I=2}$$

11.4.2 Q_4: $V-A \to V-A$, Color Mixed

Per our reasoning in section 11.3.2, we omit this. See section 11.4.1.

11.4.3 Q5: $V-A \to V+A$, Color Diagonal

This is the same as Q_3 (in our notation) (see section 11.4.1) since we are only changing the internal spin structure of the operator.

11.4.4 Q6: $V-A \to V+A$, Color Mixed

This is the same as Q_4 (see section 11.4.2) since we are only changing the internal spin structure of the operator.

11.5 Electroweak Penguin Operators

11.5.1 Q_7: $V-A \to V+A$, Color Diagonal

This is the same (in our notation) as Q_9 (see section 11.5.3), up to internal spin structure.

11.5.2 Q_8: $V-A \to V+A$, Color Mixed

This is the same (in our notation) as Q_{10} (see section 11.5.4), up to internal spin structure.

11.5.3 Q_9: $V-A \to V-A$, Color Diagonal

This operator is very similar to Q_3 (see section 11.4.1):

11.5.3.1 $I=0$

$$\begin{aligned}\left\langle 0|(\pi\pi)_{I=0}Q_9K^{0\dagger}|0\right\rangle &= \frac{3}{2}\sum_{q=u,d,s} e_q \left\langle 0|(\pi\pi)_{I=0}\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle \\ &= \frac{3}{2}\sum_{q=d,s} e_q \left\langle 0|(\pi\pi)_{I=0}\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle + \frac{3}{2}e_u Q_{1,I=0} \\ &= \frac{3}{2}\left[Q_{1,I=0}(e_u+e_d) - \frac{3}{\sqrt{3}}e_s\left(g-\frac{1}{2}(i_1+i_2)\right) + \frac{3}{\sqrt{3}}(e_d+e_s)\left(g'-\frac{1}{2}\left(i'_1+i'_2\right)\right) - \frac{3}{2\sqrt{3}}e_d\left(h'_1+h'_2\right)\right] \\ &= \frac{3}{2}\left[Q_{1,I=0}(e_u+e_d) - \frac{3}{\sqrt{3}}e_s(g-i) + \frac{3}{\sqrt{3}}(e_d+e_s)\left(g'-i'\right) - \frac{3}{\sqrt{3}}e_d h'\right]\end{aligned}$$

Distinguishing the strange quark propagator, we find

$$= \frac{3}{2}\left[Q_{1,I=0}(e_u+e_d) - \frac{3}{\sqrt{3}}e_s\left(G-G'+I'-I\right) + \frac{3}{\sqrt{3}}(e_d)\left(g'-i'-h'\right)\right]$$

11.5.3.2 $I=2$

Adding the appropriate weighting factors, we can mostly reuse our Q_3 result here, too.

$$\left\langle 0|(\pi\pi)_{I=2}Q_9K^{0\dagger}|0\right\rangle = \frac{3}{2}Q_{1,I=2}(e_u+e_d)$$

11.5.4 Q_{10}: $V-A \to V-A$, Color Mixed

Omitted. See section 11.5.3. (see section 11.3.2 for reasoning)

Chapter 12

$K \to \sigma$

We write down the process $K^0 \to \sigma$ in terms of (non-Gparity) operator expansion involving operators Q_i.

12.1 Diagram types

In general, for each term we have a three point function given in terms of times t_{dis}, t_{op}, t_{src}. t_{dis} is the distance between the four-quark operators Q_i and the σ. t_{op} is the time slice location of the four-quark operator. t_{src} is the time slice location of the K. We draw the diagrams with time flowing left to right, and arrows starting on $\overline{q}$ and ending on the q. To map these onto the contractions below, take the outer pion to also be the least inner in the bracket: $\langle 0|\pi_{\text{outer}}\pi_{\text{inner}}...|0\rangle$. We display the strange quark contraction as a double line. We assume the u,d quarks are degenerate.

We label the contractions with letters. Different topologies of the same contraction are labelled with subscripts. The strange quark propagator is denoted S^{-1} instead of the usual M^{-1} for the u,d quarks. Primed vs. unprimed diagrams are defined in sections 12.1.1 to 12.1.3.

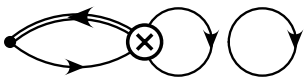
In the penguin diagrams, $V-A$ can be replaced with $V+A$ at the sink (because the sink only interacts with a gluon, photon, or Z at the operator). We know that it's the sink and not the source that gets the $V+A$ since only $V-A$ can change flavor (and we are changing strangeness in the $K \to \sigma$ process). We thus display $\pm$ in the internal operator spin structure of the sink. These diagrams should be understood to have the appropriate sign depending on the context in which they are read.

We would like to use the same notation for the diagram types of $K \to \sigma$ as $K \to \sigma$. In keeping with my convention from my previous calculation of $K \to \sigma$, we eliminate type1 diagrams entirely. We map the diagram types by contracting the pion to pion quark propagators to a point giving us the corresponding sigma diagram.

Finally, some notation needs to be defined. The 4-quark operator has two fermion loops connected to it.

Each of these loops could be a spin trace, or we could have an overall spin trace of both loops. This is the difference between the primed and unprimed diagram letters (e.g. f vs. f'). These spin traces could also be color traces, in which case $Tr(...)$ refers to both and we simply refer to the figure by the letters defined in sections 12.1.1 to 12.1.3. The other possibility is that the color trace takes the opposite (primed vs. unprimed) structure vs. the structure of the spin trace. Let x, x' be a $K \to \sigma$ diagrams. We denote the situation where the opposite trace structure should be used for the color trace by, respectively, $Tr_c(x), Tr_c(x')$.

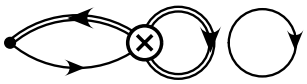
12.1.1 Type 2 Diagrams



$$\equiv g, g'$$

$$g = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\right)$$

$$g' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\right)$$

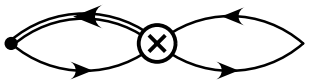


$$\equiv G, G'$$

$$g = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\right)$$

$$g' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)\cdot Tr\left(M^{-1}\right)$$

12.1.2 Type 3 Diagrams

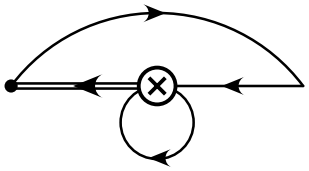


$$\equiv h, h'$$

$$h = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$

$$h' = Tr\left(M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$

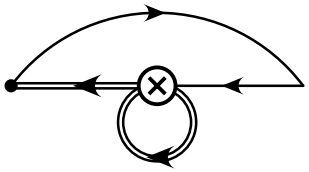
12.1.3 Type 4 Diagrams



$$\equiv i, i'$$

$$i = Tr\left(M^{-1}M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)\right)\cdot Tr\left(M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$

$$i' = Tr\left(M^{-1}M^{-1}\gamma_5 S^{-1}\gamma_\mu\left(1-\gamma_5\right)M^{-1}\gamma^\mu\left(1\pm\gamma_5\right)\right)$$



$$\equiv I, I'$$

12.2 Analysis

If we have $SU(3)$ symmetry[1],we get 12 diagrams 20 =(3: diagram types)(2: primed vs. unprimed)(2: x vs. $Tr_c(x)$). Distinguishing the strange quark propagator gives us 20.

Also, N.B. if x has $+1$ from Grassmann number commutation, x' has -1 (and vice-versa).

All correlation functions follow the general structure:

$$\langle\sigma|Q_i|K\rangle$$

12.3 Current-Current Operators

12.3.1 Q_1: Color Diagonal

$$\begin{aligned}\left\langle\sigma|(\overline{u}_\beta u_\beta)_{V-A}(\overline{s}_\alpha d_\alpha)_{V-A}|K\right\rangle &= \left\langle\sigma|(\overline{u}_\beta u_\beta)_{V-A}(\overline{s}_\alpha d_\alpha)_{V-A}\overline{d}\gamma_5 s|0\right\rangle \\ &= \left\langle\sigma|(\overline{u}_\beta u_\beta)_{V-A}(\overline{s}_\alpha d_\alpha)_{V-A}\overline{d}\gamma_5 s|0\right\rangle\end{aligned}$$

We drop the γ and $V\pm A$ factors in what follows for the sake of brevity.

$$= \left\langle\sigma|(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s|0\right\rangle$$

Per chapter 4 (where we ignore the strange quark in the scalar, i.e. $\overline{s}s$), we have the expansion for σ:

$$|\sigma\rangle = \frac{1}{\sqrt{2}}\left(u\overline{u}+d\overline{d}\right) \tag{12.1}$$

$$\Rightarrow \left\langle\sigma|(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s|0\right\rangle = \left\langle 0|\frac{1}{\sqrt{2}}\left(\overline{u}u+\overline{d}d\right)(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s|0\right\rangle$$

This gives us two terms. Let's work them out one by one.

1.

$$\frac{1}{\sqrt{2}}\left\langle 0|\overline{u}u(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s|0\right\rangle = \frac{1}{\sqrt{2}}(h-g)$$

[1]We list capital letters for the case of SU(2) symmetry, and denote the strange quark propagator by a double line.

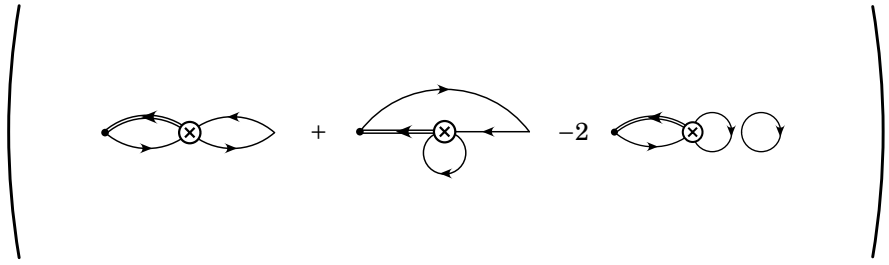
2.

$$\frac{1}{\sqrt{2}}\left\langle 0|\overline{d}d(\overline{u}_\beta u_\beta)(\overline{s}_\alpha d_\alpha)\overline{d}s|0\right\rangle = \frac{1}{\sqrt{2}}(i-g)$$

Finally, the total is

$$\left\langle 0|\sigma Q_1 K^{0\dagger}|0\right\rangle = \frac{1}{\sqrt{2}}(h+i-2g)$$

$\langle 0|\sigma Q_1 K^{0\dagger}|0\rangle =$ $\frac{1}{\sqrt{2}}$



12.3.2 Q_2: Mixed Color

Since the only flavor changing weak current is a $V-A$ current, these are the only current-current operators. We get this operator by exchanging color indices on the quarks in the 4-quark operator. Thus, any terms that are operated on by Tr_c are now non Tr_c diagrams and vice-versa. We write, trivially:

$$\left\langle 0|\sigma Q_2 K^{0\dagger}|0\right\rangle = Tr_c(Q_1)$$

The color mixed operators are always operated on by Tr_c relative to their color-diagonal counterparts (hence we will omit writing them down in sections 12.4 and 12.5).

12.4 QCD Penguin Operators

The only other distinct dimension= 4 quark operators[2] can be found by going beyond tree level. These are the penguin diagrams. Depending on whether the penguin has a gluon or a photon/Z, we get two distinct classes of diagrams. We examine the ones involving gluons in this section.

12.4.1 Q_3: $V-A \to V-A$, Color Diagonal

$$\begin{aligned}\left\langle 0|\sigma Q_3 K^{0\dagger}|0\right\rangle &= \sum_{q=u,d,s}\left\langle 0|\sigma\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle \\ &= \sum_{q=d,s}\left\langle 0|\sigma\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle + Q_1\end{aligned}$$

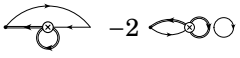
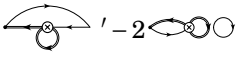
Let's calculate each of the two terms separately.

We get all of Q_1 because σ is symmetric with respect to d,u interchange (any contraction involving $u,\overline{u}$ from the operator and a pion now involves $d,\overline{d}$ from the operator, and a different term in σ with $u \leftrightarrow d$).

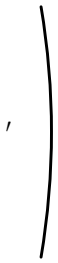
We also get additional terms from contractions of the operator with the K/itself. These contractions weren't available when the operator had a $\overline{u}u$ (as in sections 12.3.1 and 12.3.2). One possibility is the q from the $\overline{q}q$ contracting with K, and the other possiblity is the $\overline{q}q$ contracting with the $\overline{s}d$ from the operator.

The s term must have a loop on the operator. Therefore, all terms from Q_1 that don't have this loop are excluded. This gives us finally,

$$\left\langle 0|\sigma Q_3 K^{0\dagger}|0\right\rangle = 2Q_1 - Q_1' + \frac{1}{\sqrt{2}}\left((I-2G)-(I'-2G')\right)$$

$\left\langle 0|\sigma Q_3 K^{0\dagger}|0\right\rangle = 2Q_1 - Q_1' +$ $\frac{1}{\sqrt{2}}\Bigg($ +  $\Bigg)$ $-\frac{1}{\sqrt{2}}\Bigg($ +  $'$ $\Bigg)$

[2]To see these exhaust the available operators, consider that one of our two fermion lines (meaning a $\overline{q}q$ in the operator's quark content) is determined by needing to change strangeness. Then, either the other fermion line makes contact with a W, or it doesn't. Suppose it does. If this W originated on the distinguished fermion line, we fix the flavors of the other line, and get a current-current contribution. If it originated on the undistinguished line, then it doesn't do anything to distinguish this line's flavors. Also, there can be no initial to final state flavor change. Now, if this line doesn't make contact with a W, it must contact the operator through a Z,g,γ. We thus arrive at the penguin operators.



12.4.2 Q_4: $V-A \to V-A$, Color Mixed

Per our reasoning in section 12.3.2, we omit this. See section 12.4.1.

12.4.3 Q5: $V-A \to V+A$, Color Diagonal

This is the same as Q_3 (in our notation) (see section 12.4.1) since we are only changing the internal spin structure of the operator.

12.4.4 Q6: $V-A \to V+A$, Color Mixed

This is the same as Q_4 (see section 12.4.2) since we are only changing the internal spin structure of the operator.

12.5 Electroweak Penguin Operators

12.5.1 Q_7: $V-A \to V+A$, Color Diagonal

This is the same (in our notation) as Q_9 (see section 12.5.3), up to internal spin structure.

12.5.2 Q_8: $V-A \to V+A$, Color Mixed

This is the same (in our notation) as Q_{10} (see section 12.5.4), up to internal spin structure.

12.5.3 Q_9: $V-A \to V-A$, Color Diagonal

This operator is very similar to Q_3 (see section 12.4.1):

$$\begin{aligned}\left\langle 0|\sigma Q_9 K^{0\dagger}|0\right\rangle &= \frac{3}{2}\sum_{q=u,d,s} e_q \left\langle 0|\sigma\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle \\ &= \frac{3}{2}\sum_{q=d,s} e_q \left\langle 0|\sigma\overline{q}_\beta q_\beta \overline{s}_\alpha d_\alpha \overline{d}s|0\right\rangle + \frac{3}{2}e_u Q_1 \\ &= \frac{3}{2}\left[Q_1(e_u+e_d) - Q_1' e_d + \frac{1}{\sqrt{2}}e_s\left((I-2G)-(I'-2G')\right)\right]\end{aligned}$$

12.5.4 Q_{10}: $V-A \to V-A$, Color Mixed

Omitted. See section 12.5.3. (see section 12.3.2 for reasoning)

12.6 $K \to \pi$

This study is well beyond the scope of this work, but we can roughly use chapter 12 with a few important exceptions. First, the operator vertex must come with even parity, so all contributions which have an odd number of γ_5 must vanish. Furthermore, we know there are no disconnected diagrams if isospin is a good symmetry. Beyond this, we know eq. (12.1) must come with a relative minus sign when it becomes a π (specifically, π^0 for charge conservation). The subsequent contractions are left as an exercise.

Chapter 13

Conclusions/Future Work

We first list open problems associated with $\pi\pi$.

13.1 Technical List of Future Work

13.1.1 Unsolved

1. More optimized radii for sigma, rho, pion

2. Why are the burst buffer reads so slow? (slowness of i/o)

3. Additional operators, point source, etc.

4. More hits for specific meson fields. This study is under way at present, but will not be concluded before the completion of this thesis.

5. Why is the error on the scattering length so much smaller in Lüscher's expansion method?

6. Long term plan for merging A2A library with Grid

7. Around the world systematic error (explicitly measure)

8. Why is does γ_5 herm reduce errors on strange prop in k->pipi?

9. Why is C. Lehner's MADWF code faster (no split-CG)?

13.1.2 Under-solved

1. Comms: Broadcast, split v. unsplit communicator, threads vs. processes. Intel results aimed at fixing KNL comms via multi-threading are still in process at the time of this writing, perhaps tuning of algorithms is necessary (via, e.g., mpitune)?

2. Tuning of contraction code for 1 ppn (an attempt was made, but so far this tuning did not improve on 4 ppn).

3. $I = 1$ tension with C. Lehner and A. Meyer. Maybe we need a new (vector) operator in this channel?

13.1.3 Unimplemented

1. Distillation for k->pipi

2. Kaon point source (consult prior g-parity work of C. Kelly)

3. Free field (non-interacting) check of the draft periodic $K \to \pi\pi$ code.

4. Coarse deflation. Deflation normally takes 5 min per group of solves on 24^3. C. Lehner has code which uses the compressed evecs to deflate the solve. This could be quite cumbersome to implement, however, given the already cumbersome eigenvector interface. CG on 32^3 with 4000 evecs may require such work, however. Sergey suggests we might do a less rudimentary cache blocking which might also speed things up.

5. Cross terms in the a2a (from C. Lehner, and then subsequent discussion with L. Jin). If we think about modes as being either high (all-low) or low (evecs), then cross terms (high-low, low-high) arise over the two indices of a meson field. If a meson field were pure low or pure high, then we would have (probably) a large correlated term (low-low) and a small correlated term (high-high) . The result could be fixed up with AMA as necessary. Others (at Jlab) are doing this, apparently, with a similar block diagonal structure. No other details have yet emerged, but it may be worth it to wait to see what they do. I'm fairly sure the code I've written for this will work, but it still requires implementing a high level loop over the mode blocks of contractions. We can correct any systematic error introduced via AMA.

6. γ^5 hermiticity to reduce contractions

7. Block solver for additional hits

8. Simultaneous fits (non-GEVP fits to GEVP matrices).

9. Display interaction energy on plots instead of total (config switch)
10. Multi-exponential fits on the gevp.
11. Phase shift code for energies below $\pi\pi$ threshold.
12. Scattering length for $I = 0$.
13. Take the continuum limit.
14. $I = 1$ moving frame phase shift code.

13.1.4 Under-implemented

1. statistics on 24^3
2. statistics on 32^3
3. 32^3 evecs (1.4, 1 GeV ensembles. C. Jung has done quite a few on 1.4, how many do we still need?)
4. Fully utilize auxiliary symmetry. There remain about 10% of momentum combinations which can be eliminated, but the code could become more complicated. We may need to make things more abstract first.

13.1.4.1 Things to test

1. My data processing/fitting code (for $I = 0$ especially).
2. different radii for sigma and rho
3. $K \to \pi\pi$ The main calculation. How general is the G-parity code? How much is left to do?
4. irrep cross terms in moving frames
5. Comparison to Gparity.
6. Different block parameters for contraction multiplies (tuning for 1 ppn, e.g.)

13.1.5 Not solved, but low priority

1. A totally split madwf. Several barriers prevent progress: memory (primarily); how to deflate with split sources
2. Parity symmetry unexploited (we are computing $I = 1$)

3. Axis symmetry unexploited (we want the building blocks of irreps. We may average over rotations later.)

4. Nested threading of meson field block multiply (is this even feasible?)

13.2 Conclusion

We set out to calculate $\pi\pi$ elastic scattering phase shifts from lattice QCD with physical quark mass. We obtain fair to good agreement with phenomenology, but more work needs to be done to understand the remaining systematics. These include interacting around the world terms, fit range dependence, excited state contamination (especially for $I = 1$), and the transition amplitude from two-pion to four-pion states. A full systematic error budget is necessary before we can reach complete conclusions of $\pi\pi$ phase shifts.

Additionally, this initial study had fairly low statistics, and the fit range selection was not optimized to minimize errors. Our production of more hits on the 24^3 ensemble will soon be complete, as well as our analysis of $I = 1$ moving frames, scattering length determinations, and the formal procedure of taking a continuum limit. We thus should conclude that the phase shifts determined here are likely to soon be substantially improved. We also should acknowledge the other physical point $\pi\pi$ studies currently under way, including a study using distillation and another using A2A g-parity boundary conditions. A joint publication of all three studies is planned, which should have finalized phase shift and energies by the end of the year (again, as of this writing).

While much groundwork has been laid for periodic $K \to \pi\pi$, more work is needed to test the code before we are ready for production. We also plan to test numerically the various new GEVP developments to assess their potential for improving lattice spectroscopy generally. Also, while this study is a good start in terms of learning about the $\pi\pi$ system, more rigorous testing needs to be done to asses alternative models and approaches. The many compatible p-values we obtain when fit range averaging may indicate that we do not fully understand the range of possible parameters (and so underestimate the systematic error).

Finally, we should acknowledge that substantial software/hardware development time has passed since the beginning of the study, so further work could benefit from a critical reassessment of the production tools. Looking to the future, we might plan to port the code fully to Grid, write a GPU version of the code, and/or plan for studies at the charm scale. For hadronic physics, studies of the $\pi\pi$ system will likely be useful for testing and pedagogical purposes for years to come.

Appendices

Appendix A

GEVP Effective Mass Plots

How to read the plots:

There are a few items to note. First, there are two kinds of fits to the effective mass displayed here. One uses a systematic error ansatz (see section 9.7), while the other fits to a constant. The energies and phase shifts are p-value weighted averages, in general. However, for the fits of section 9.7, most of the time it was difficult to obtain fit ranges with good p-values, so many are fits to a single fit range. The energies displayed on the plot are the p-value weighted averages. The fit range actually plotted is chosen to be a representative which minimizes the maximum difference between energies, and the χ^2/dof displayed is that of the representative range. If a χ^2/dof is not displayed, this is because it was > 2. However, the average p-value was always above the cutoff (of 0.1). Overfitting is allowed only if all other fit ranges in the window did not give acceptable χ^2.

The grey bands in the constant fits represent the $\pm$ error on the central fit value (and this tolerance is the p-value weighted average). The fits of section 9.7 do not display a tolerance. Sometimes, time slices are skipped in a given fit range. This is reflected on which points have a grey band around them. In general, there is also a fit line which goes through the middle of the grey band (although this is often invisible).

The single lines displayed as "disp" are the free energies displayed from the dispersion relation. All energies are given in lattice units. "Exact" indicates that we have performed the necessary AMA correction (see section 2.6). $t - t_0 =< integer >$ is the distance between GEVP LHS and RHS times, which we fix to be a constant. In, e.g., $185, 7$, the first number is the total number of configs, while the second is the number of exact configs used to correct the sloppy samples. The $\vec{p}_{cm}$ indicated includes all permutations and negative versions of this three-momentum.

The isospin (or some indication of the isospin like inclusion of the σ) is displayed as, e.g., $I2$.

Matrix subtraction is indicated as "matdt<integer>." The integer indicates the δ_t used in matrix subtrac-

tion (see eq. (8.22) in section 8.4.1).

Fit ranges were in general set to be arithematic sequences. However, the late time cut (where we cut out points with large error bars; see item 5 in section 9.6.2) is applied afterwards, so not all the sequences displayed have a constant difference between included points.

A.1 $I = 0$ GEVP Plots

A.1.1 24^3 DSDR

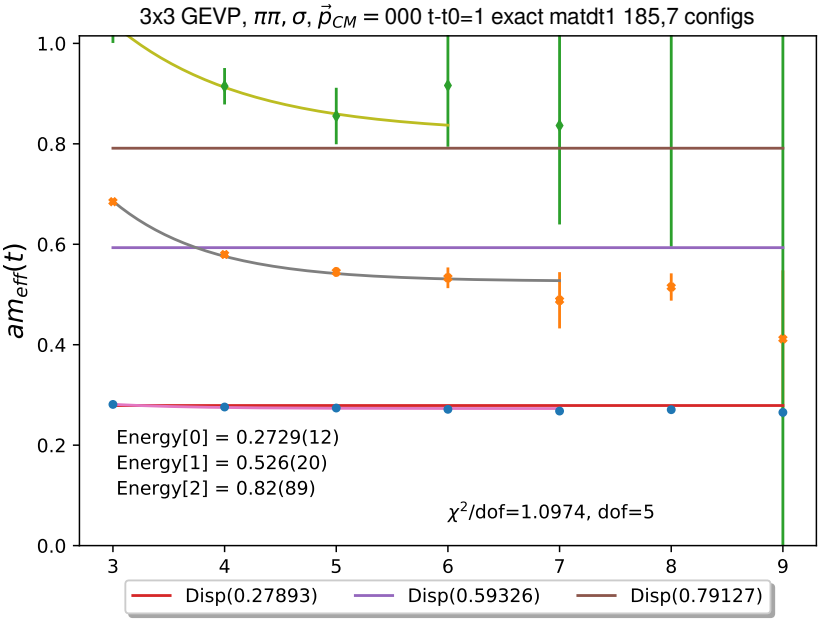


Figure A.1: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
277.0(1.2)	N/A, but scattering length (lattice units) = −24.81(22)
534(21)	48(13)
833(908)	-19(66)

Table A.1: p0 Numerical Values, the ground state is below $2m_\pi$ in energy, so the phase shift (if it is defined via analytic continuation) is complex (and not shown)

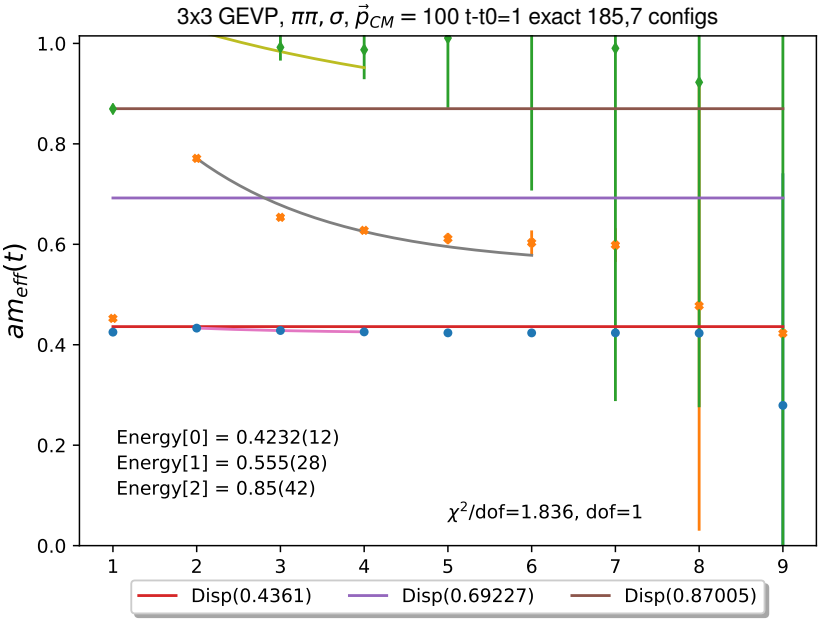


Figure A.2: p1

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
337.5(1.5)	10.77(99)
497(32)	-82(18)
819(447)	13(562)

Table A.2: p1 Numerical Values

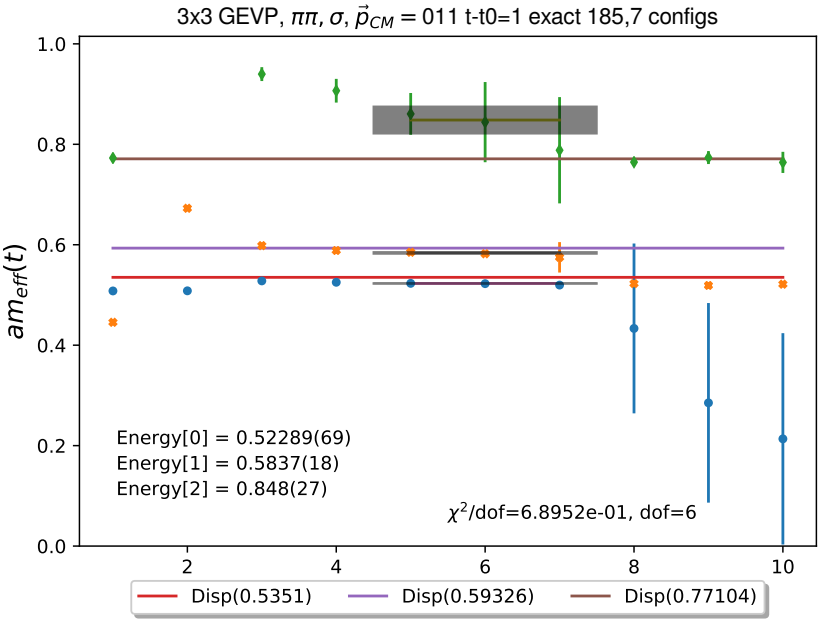


Figure A.3: p11

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
374.8(1.0)	14.68(80)
458.0(2.3)	30.9(7.8)
775(30)	-80(381)

Table A.3: p11 Numerical Values

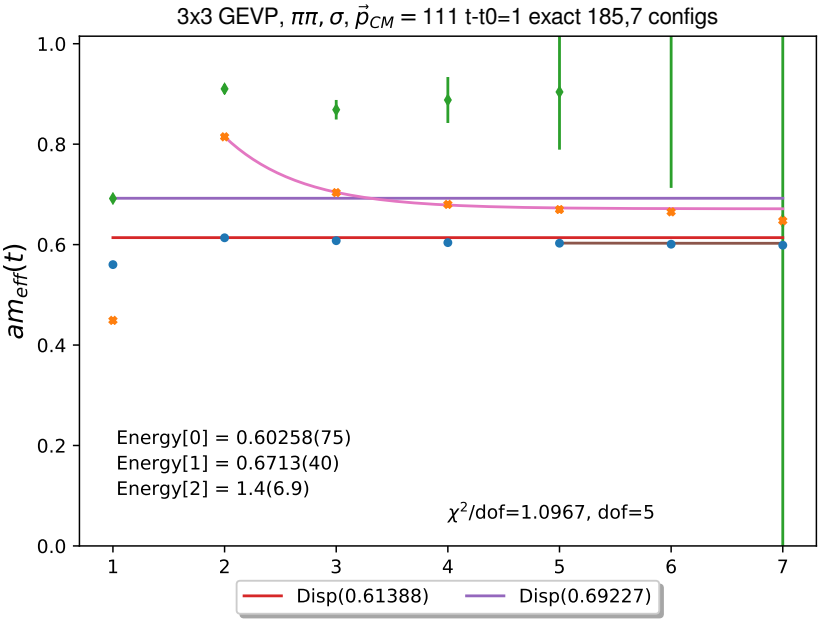


Figure A.4: p111

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
402.8(1.2)	16.5(2.7)
502.4(5.5)	28.4(8.1)
1347(7417)	50(329)

Table A.4: p111 Numerical Values

A.1.2 32ID-Fine

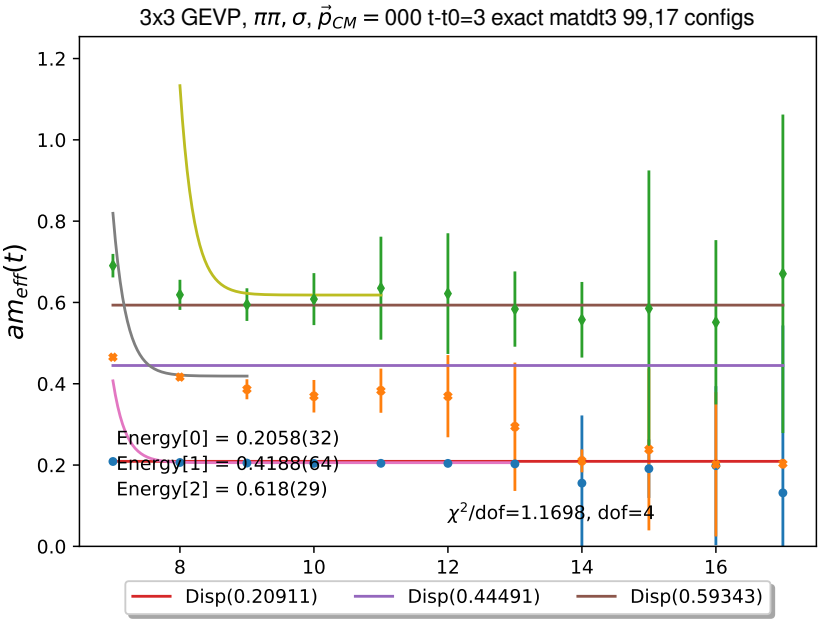


Figure A.5: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
283.7(4.3)	N/A, but scattering length (lattice units) = −8.1(1.9)
577.3(8.8)	26.6(6.0)
852(40)	-25(29)

Table A.5: p0 Numerical Values; the ground state is below $2m_\pi$ in energy, so the phase shift (if it is defined via analytic continuation) is complex (and not shown)

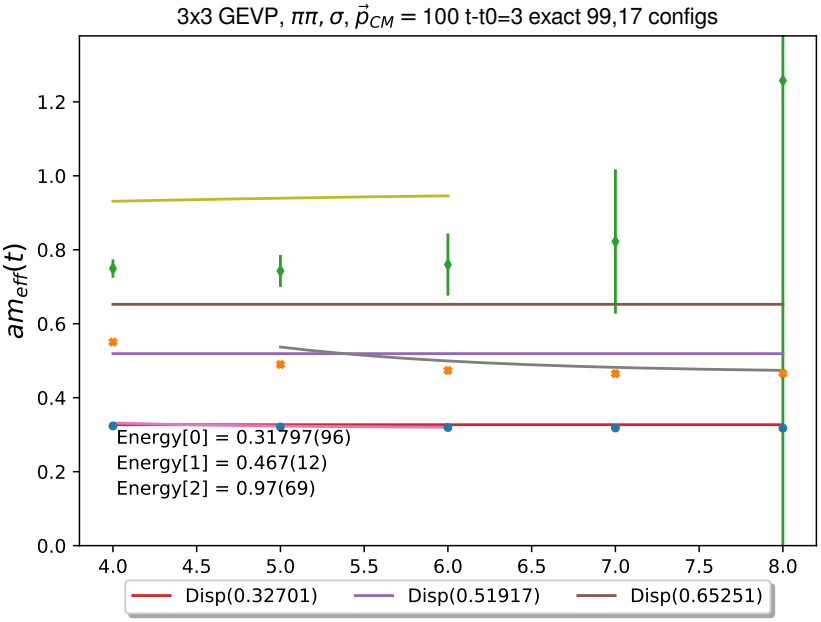


Figure A.6: p1

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
344.7(1.7)	10.0(1.0)
584(18)	51(10)
1304(976)	-20(225)

Table A.6: p1 Numerical Values

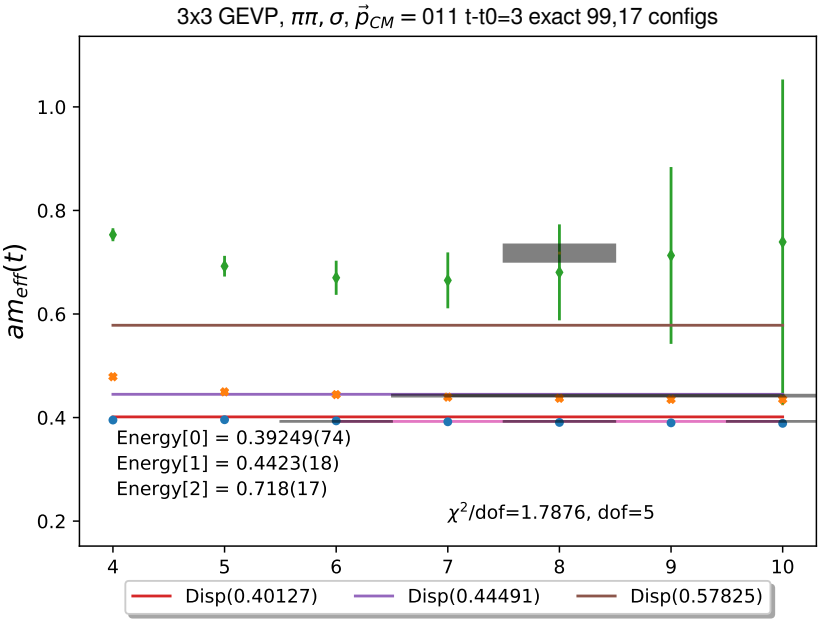


Figure A.7: p11

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
380.3(2.2)	15.5(1.6)
465.8(8.0)	35(26)
994(51)	-16(76)

Table A.7: p11 Numerical Values

A.2 $I = 1$ GEVP Plots

A.2.1 24^3 DSDR

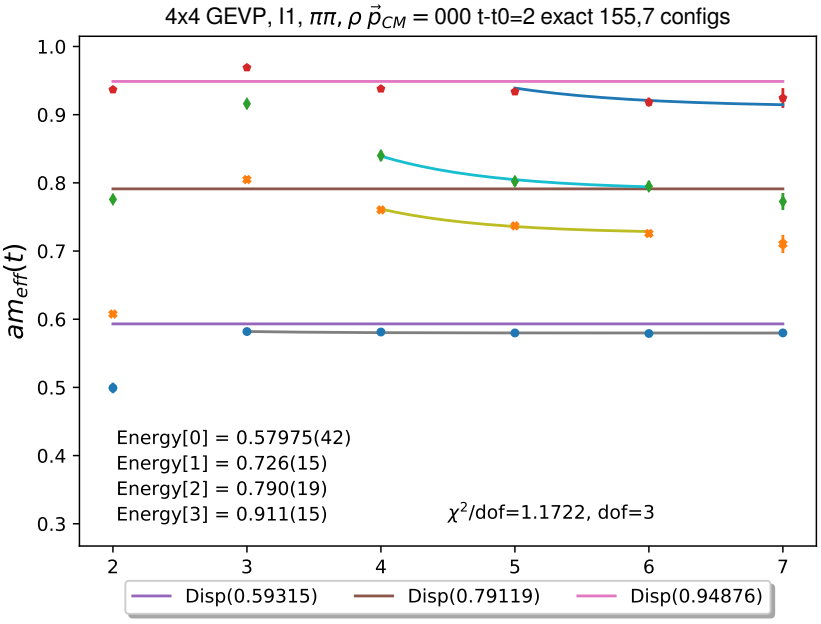


Figure A.8: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
588.44(42)	10.93(32)
737(16)	54(14)
801(20)	2(15)
925(15)	65(23)

Table A.8: p0 Numerical Values

A.2.2 32ID-Fine

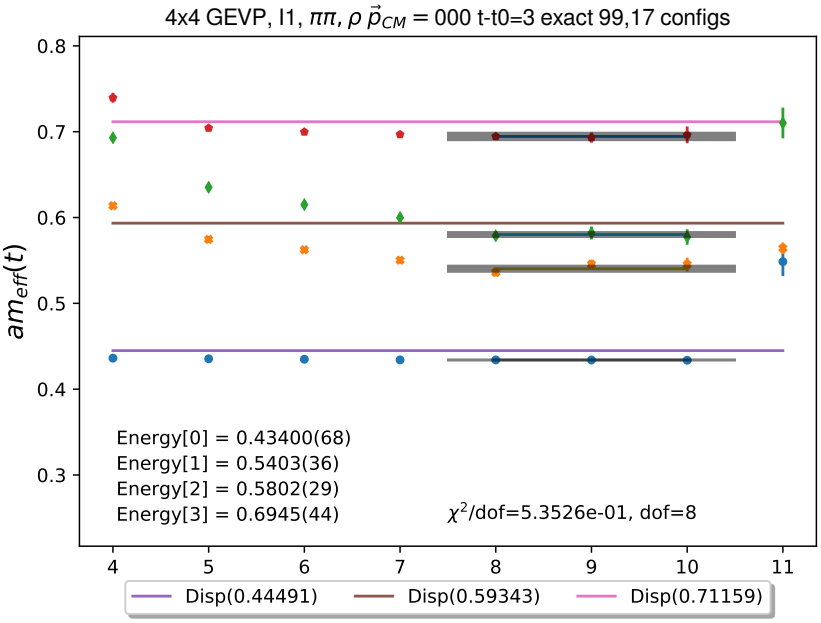


Figure A.9: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
598.23(93)	11.64(69)
744.8(5.0)	58.7(4.4)
799.7(4.1)	193.8(3.1)
957.3(6.1)	211(10)

Table A.9: p0 Numerical Values

A.3 $I = 2$ GEVP Plots

A.3.1 24^3 DSDR

todo: p0 plot

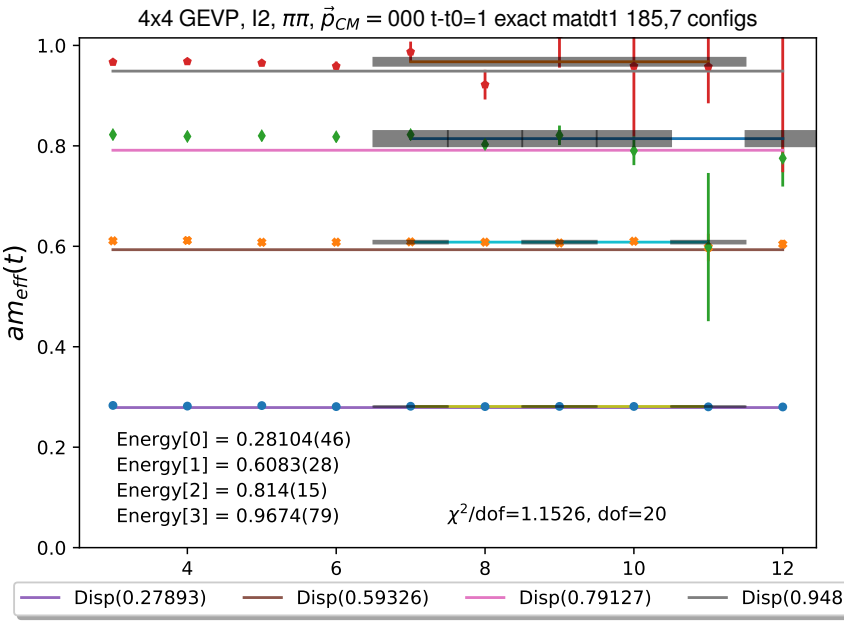


Figure A.10: p0

Table A.10: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
285.26(47)	-0.196(83)
617.4(2.8)	-12.43(62)
827(15)	-17.6(2.9)
982.0(8.0)	-28.5(9.2)

p0

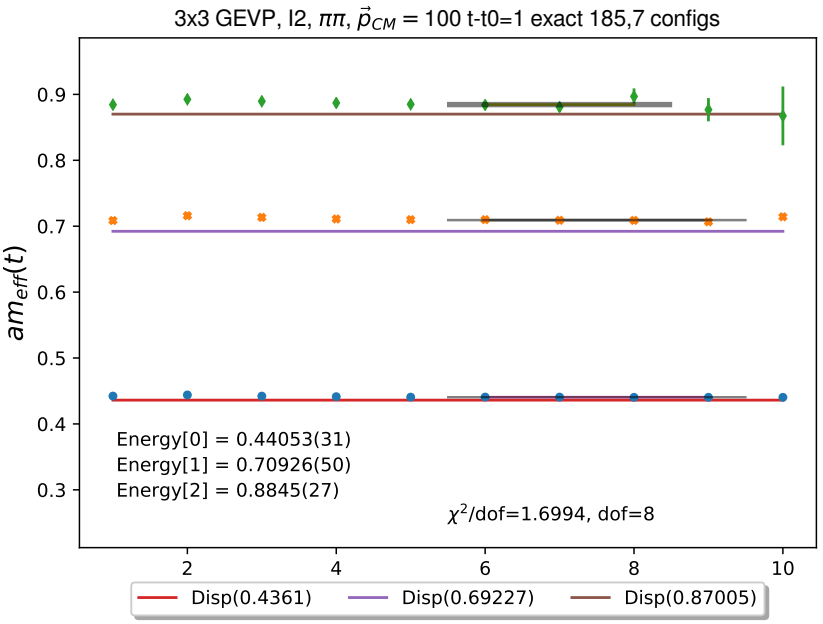


Figure A.11: p1

Table A.11: p1

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
359.61(39)	-3.12(23)
669.06(55)	-14.39(54)
857.6(2.9)	-18.3(4.8)

p1

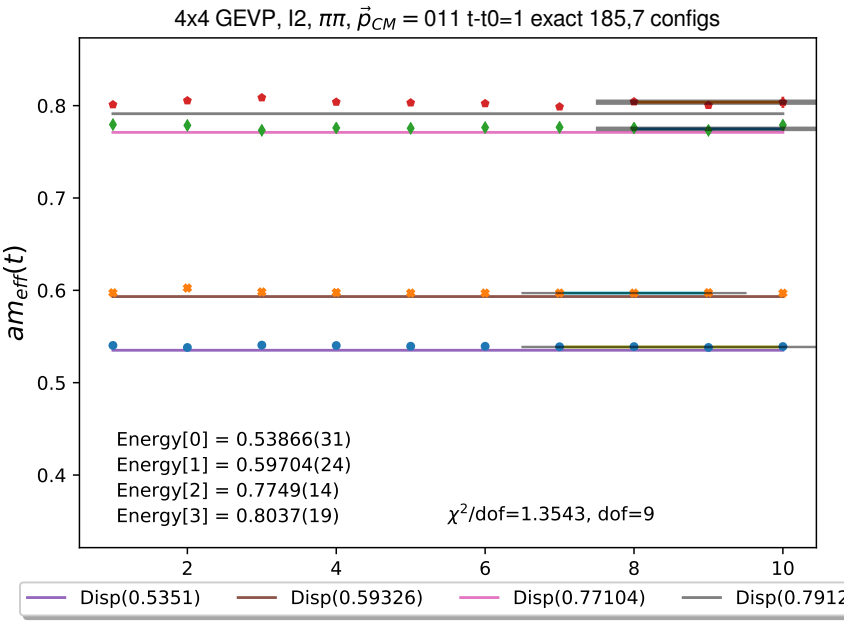


Figure A.12: p11

Table A.12: p11

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
397.12(43)	-3.88(38)
475.40(31)	-7.09(45)
690.9(1.7)	-9.1(4.8)
724.1(2.1)	-20.6(2.5)

p11

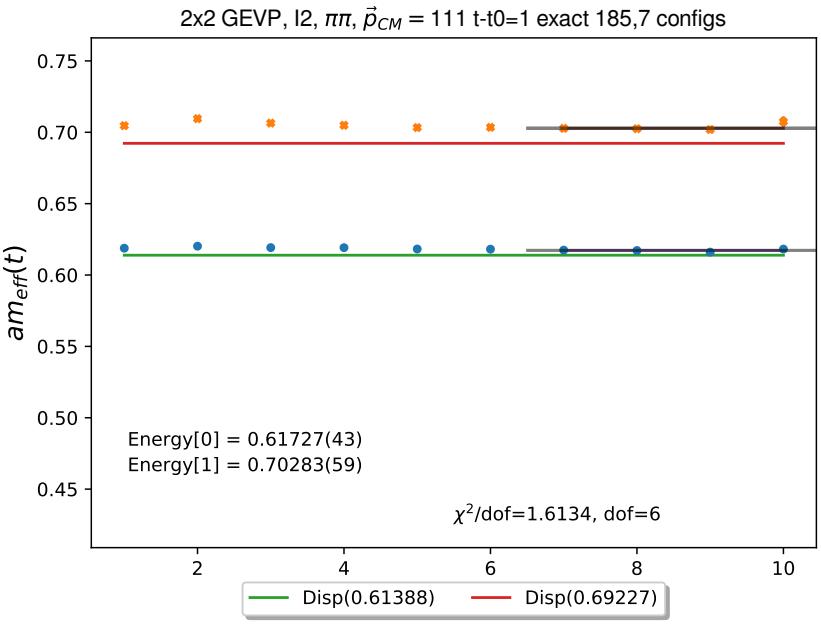


Figure A.13: p111

Table A.13: p111

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
425.10(65)	-4.81(73)
545.04(78)	-9.97(39)

p111

A.3.2 32ID-Fine

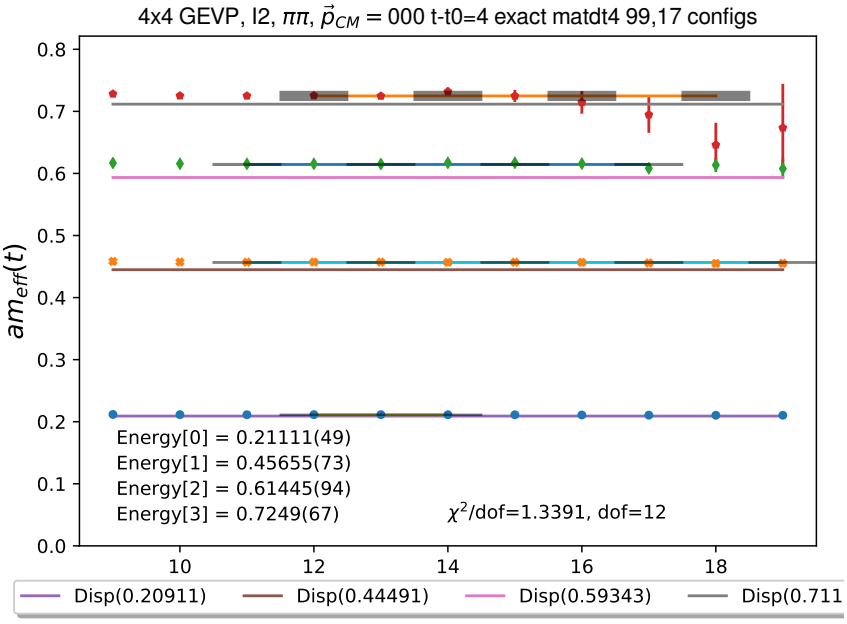


Figure A.14: p0

Table A.14: p0

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
291.00(68)	-0.35(14)
629.3(1.0)	-12.94(96)
847.0(1.3)	-21.4(1.6)
999.2(9.3)	-27.3(9.2)

p0

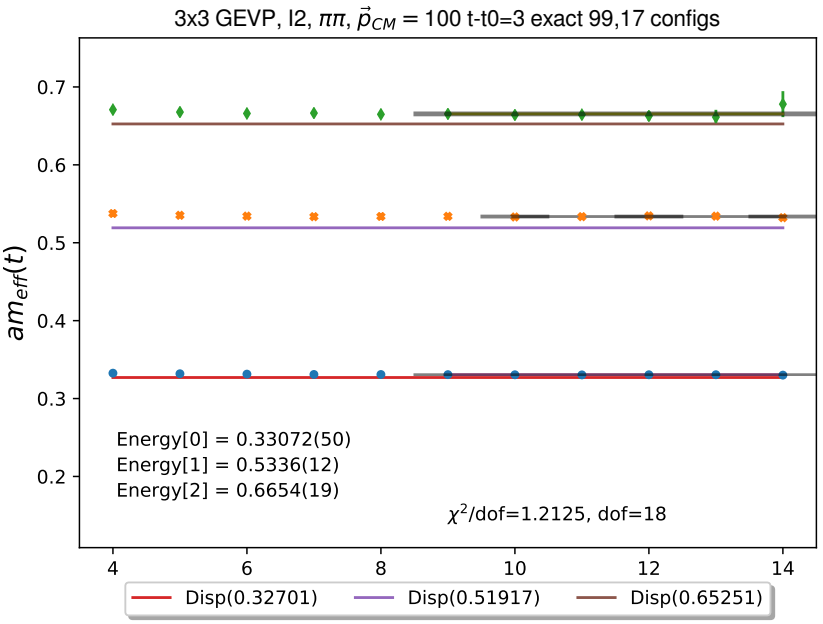


Figure A.15: p1

Table A.15: p1

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
366.83(85)	-3.62(45)
684.0(1.9)	-16.5(1.4)
876.3(2.7)	-21.6(5.1)

p1

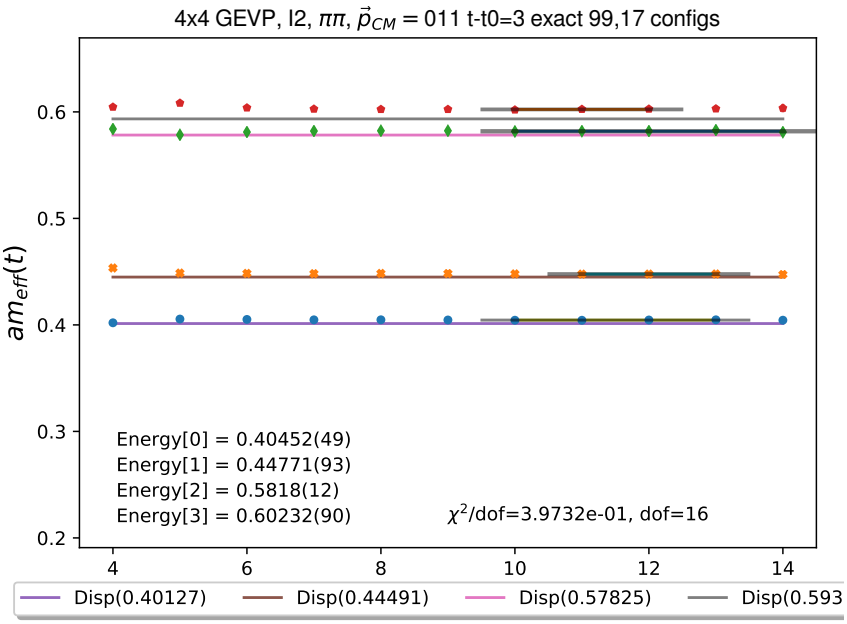


Figure A.16: p11

Table A.16: p11

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
405.48(93)	-5.01(88)
484.1(1.6)	-7.2(2.2)
704.7(1.8)	-12.7(5.6)
736.7(1.4)	-19.8(1.9)

p11

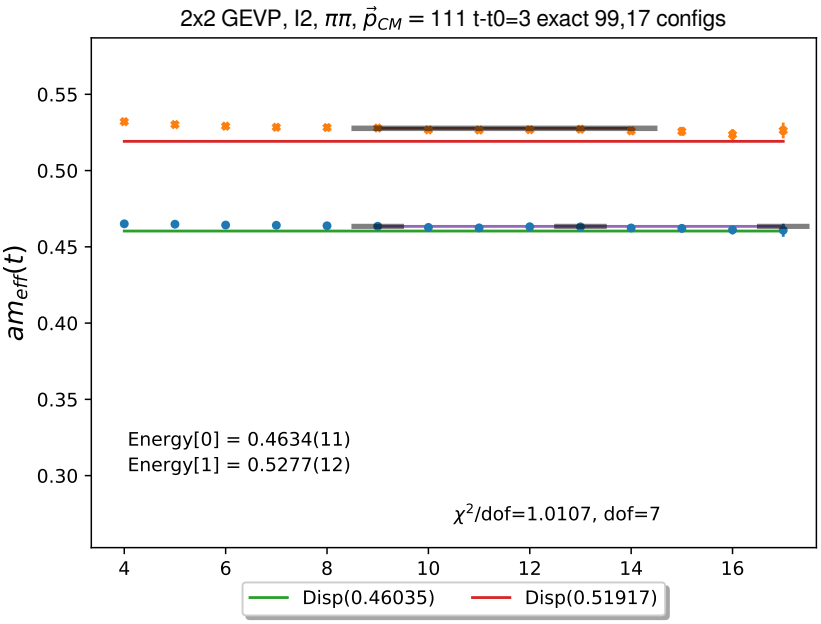


Figure A.17: p111

Table A.17: p111

$\sqrt{s}$ (MeV)	Phase Shift (degrees)
433.9(2.2)	-6.2(1.6)
556.2(2.1)	-10.8(1.5)

p111

Appendix B

Contraction Code Consistency Check

We perform a statistical comparison of the all-to-all periodic boundary condition calculation of pipi scattering from Daiqian Zhang's thesis ([30]). We perform the check on the KNL cluster at Brookhaven National Laboratory using Grid and updated A2A code. Specifically, we look at $I = 0,2$ pipi two-point correlation function as well as the single pion correlation function on an ensemble of 16^3 lattices. The comparison uses contraction data from [30] as well as data files from our runs, performs jackknife resampling, and derives effective mass plots. Good agreement is found between the resulting figures, as well as to the Daiqian's numbers quoted in his individual posting[32].

B.1 A2A Parameters

Details of our measurement setup and the setup of [30] can be found in section 3.1.

B.2 Contractions

Clebsch-Gordan coefficients tell us how to project onto definite isospin. We sum over topologies as opposed to average over them, so the coefficients we obtain are not the same as [7], but are consistent. Also, the diagramatic sum treats as equivalent diagrams that are the same contraction but the arrows are reversed. Each diagram class is separated by parentheses.

The diagramatic results for $I = 0$:

(omitted, see section 4.3.3)

The diagrams in parentheses are a specific diagram class, and the sum inside alludes to the sum over topologies which occurs in the generation (non-analysis) phase of the calculation. If we sum over topologies,

we get, in the notation of [7]:

$$\langle \pi\pi|\pi\pi\rangle_{I=0} = 3V + D - \frac{3}{2}R + \frac{1}{2}C$$

Averaging over topologies gives us

$$\langle \pi\pi|\pi\pi\rangle_{I=0} = 3V + 2D - 6R + C$$

We do not perform the typical vacuum subtractions as in Qi's thesis (e.g.):

$$V(t) = \frac{1}{L_t}\sum_{t_{src}} \left(\left\langle O(t+t_{src})O^\dagger(t_{src}) \right\rangle - \langle O(t+t_{src})\rangle \left\langle O^\dagger(t_{src}) \right\rangle \right)$$

where the bracket denotes averaging over configurations. Instead, we enforce the last term's ($\langle O\rangle \langle O^\dagger\rangle$) time translational invariance. This means that we can absorb the vacuum subtraction into the overall constant which occurs in two particle two point correlation functions (occurring when one particle wraps around the periodic box and one particle does not, but they both annihilate on the same time slice).

For $I = 2$:

B.2.1 $I = 2$

(diagram sum is again omitted, see section 4.4.4)

sum over topologies:

$$\langle \pi\pi|\pi\pi\rangle_{I=2} = D - C$$

average over topologies:

$$\langle \pi\pi|\pi\pi\rangle_{I=2} = 2(D - C)$$

B.2.2 Single Pion Two-Point Correlation Function

The single pion correlation function is automatically projected onto $I = 1$ (not pictured):

$$Corr(t) = \frac{1}{L_t} \sum_{t_{src}} \left\langle \pi(t + t_{src}) \pi^{\dagger}(t_{tsrc}) \right\rangle$$

B.3 Results

B.3.1 Effective Mass Plots

See below for graphs comparing pipi $I = 0, 2$ as well as the single pion correlation function.

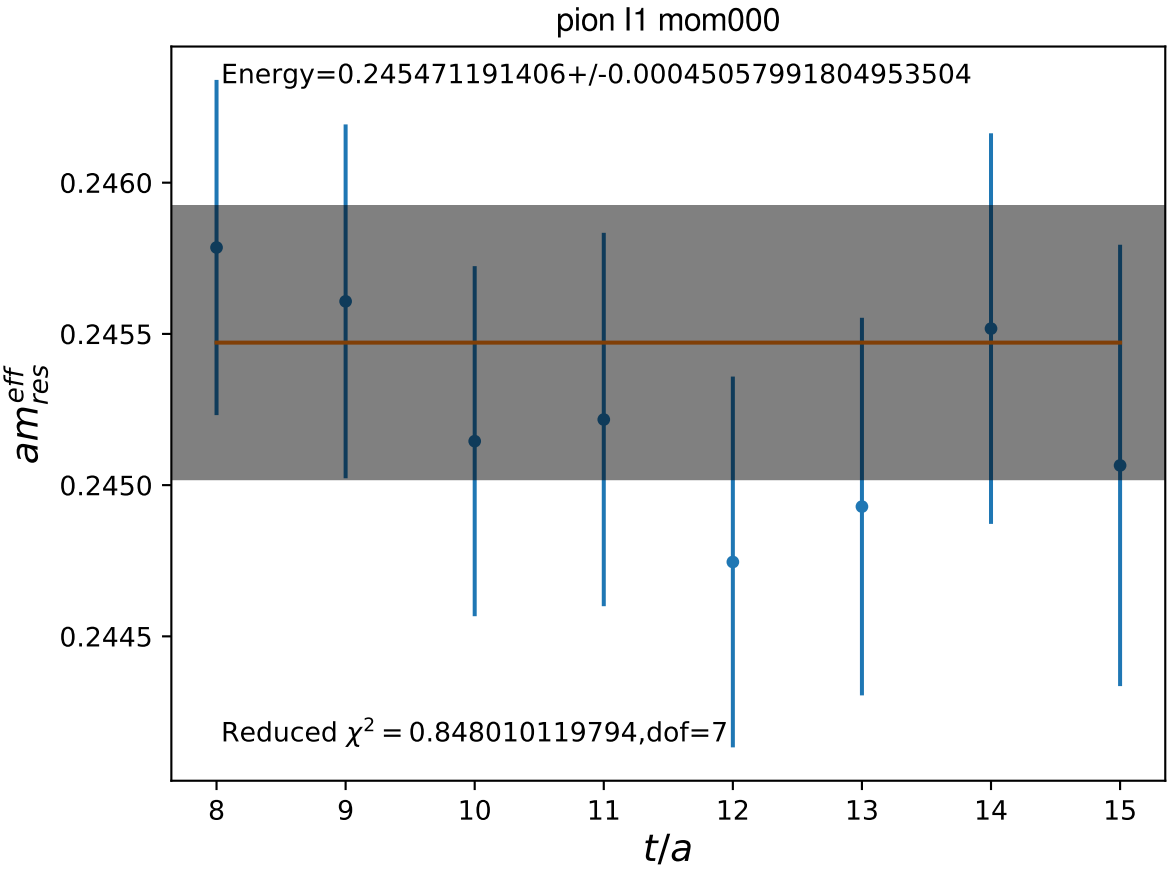
A few notes on the figures:

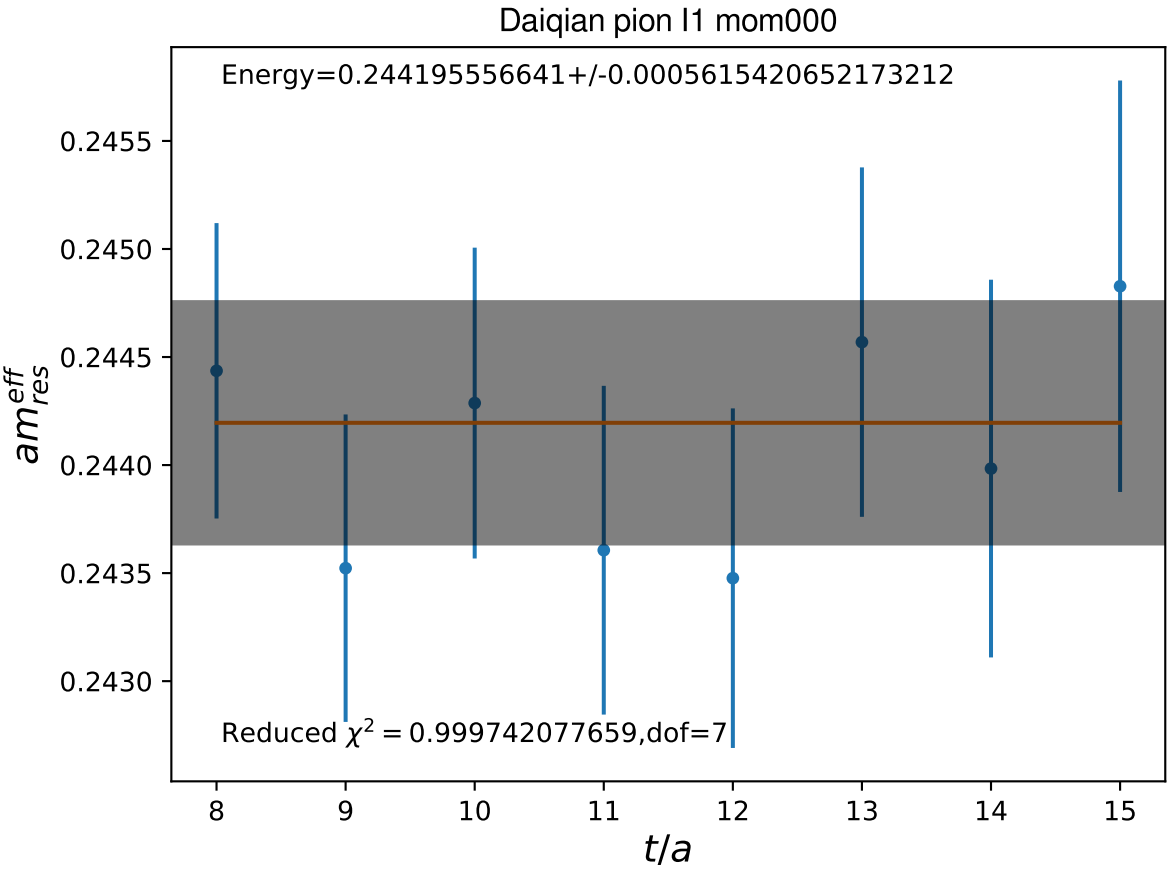
1. A shaded box is displayed which is centered on the fitted effective mass which is displayed as "Energy = ... +/- (err in Energy)"

2. When the fit curve is a constant, the fitted points are from $arccosh\left(\frac{C(t)+C(t+2)-2*D}{2(C(t+1)-D)}\right)$ where $C(t)$ is the correlation function at t. D is human-guessed parameter subtracted from the correlation function to account for the wrap-around contribution of one pion going forwards in time and one going backwards annihilating on the same time slice. D is counted as a fit parameter in the dof (although it's not really an un-biased fit parameter, and as its use is not really justifiable).

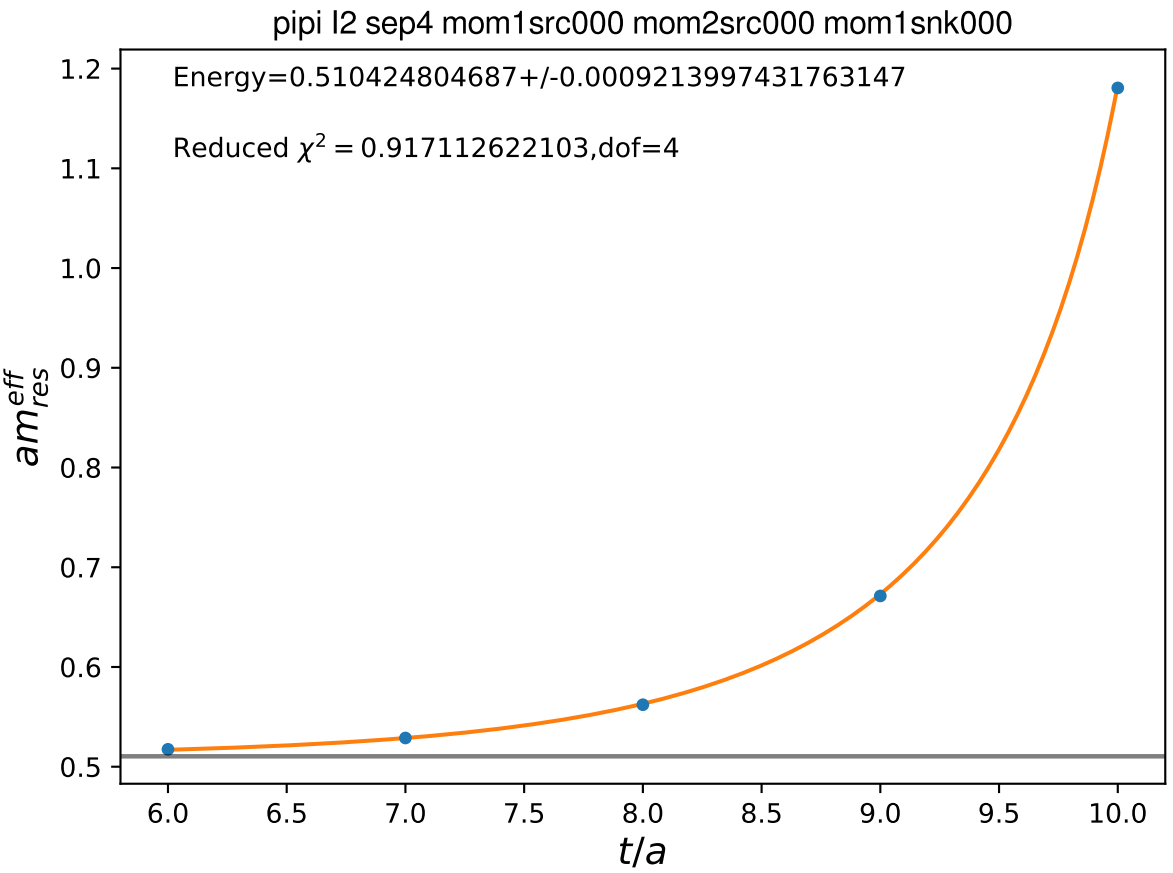
3. The ascending the curves are one parameter cosh fits where the points plotted are $log\left(\frac{C(t+1)-C(t)}{C(t+2)-C(t+1)}\right)$. This curve will be mostly flat (feeling a single exponential) until it feels the full cosh in the middle.

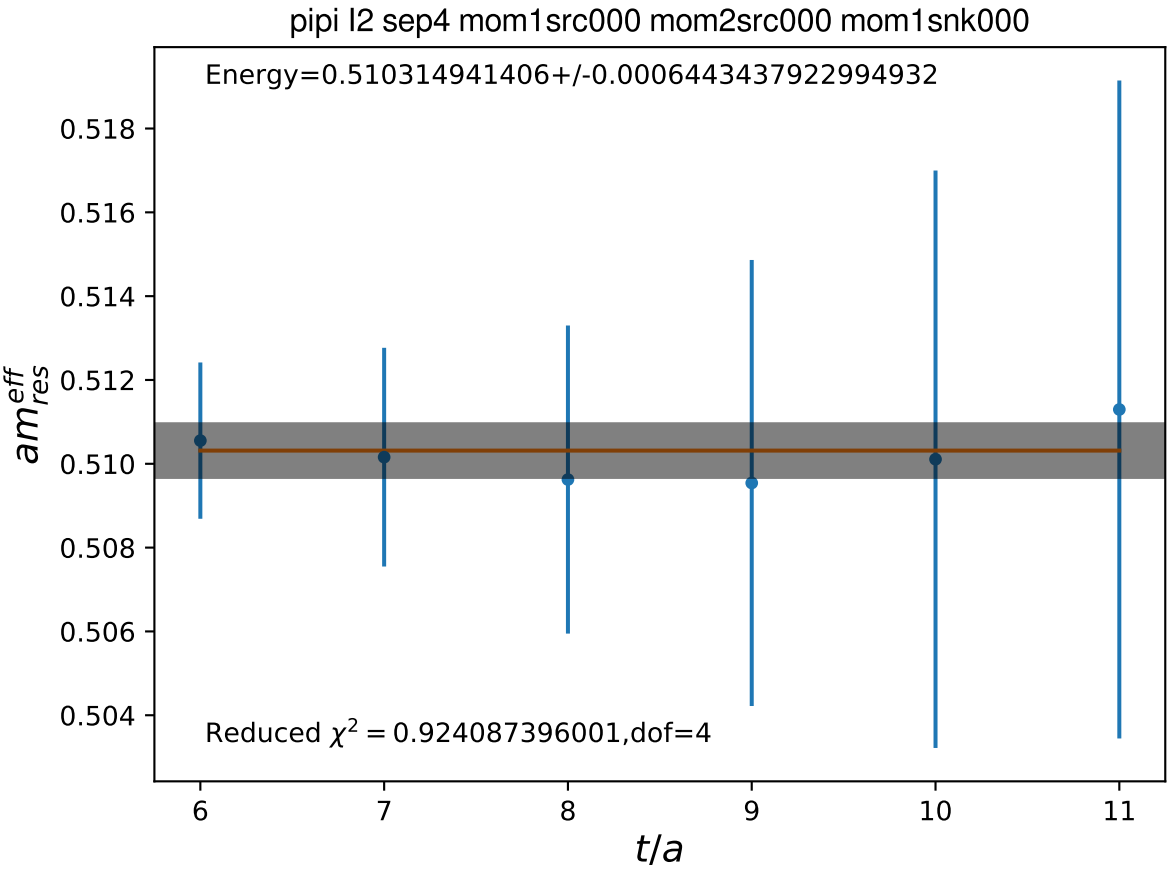
B.3.1.1 Single Pion

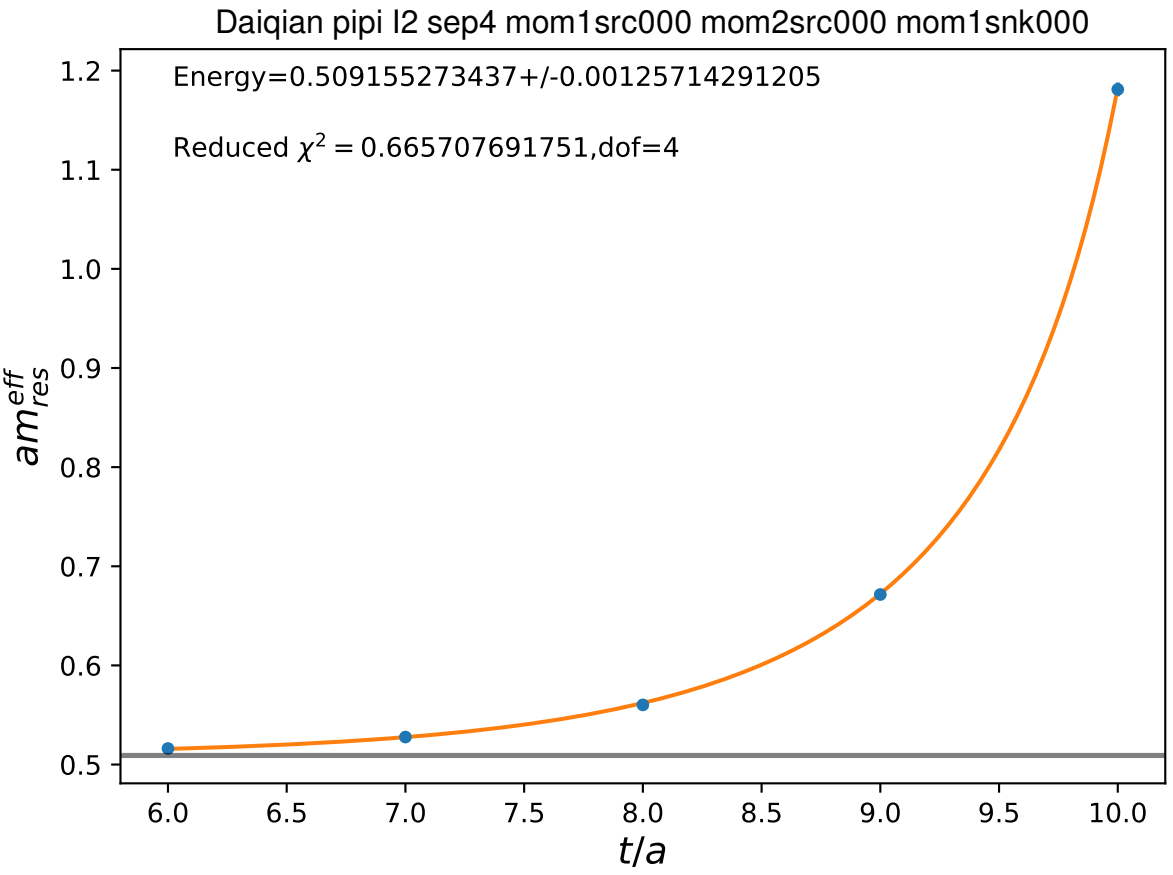


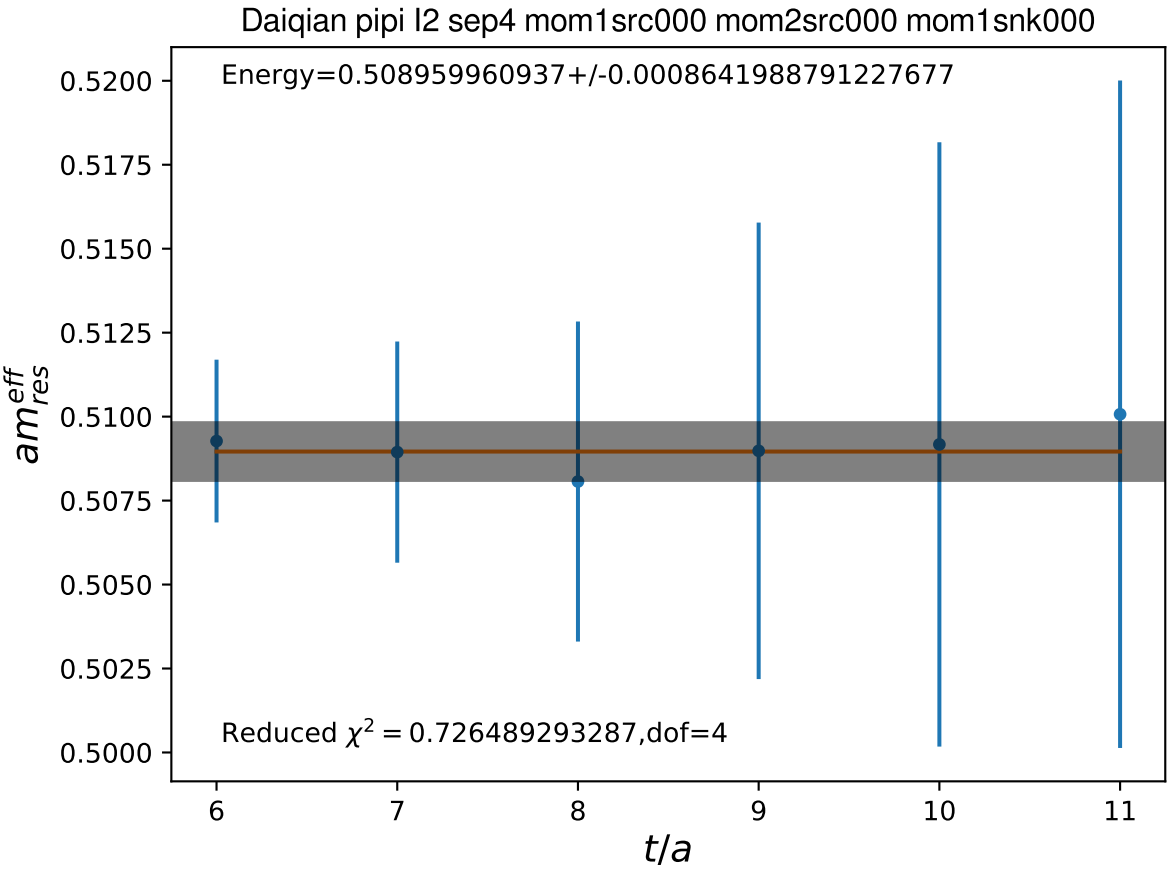

Daiqian pion I1 mom000
Energy=0.244195556641+/-0.0005615420652173212
0.2455
0.2450
0.2445
0.2440
0.2435
0.2430
am_{res}^{eff}
Reduced $\chi^2 = 0.999742077659$,dof=7
8
9
10
11
12
13
14
15
t/a

B.3.1.2 Pipi, $I = 2$

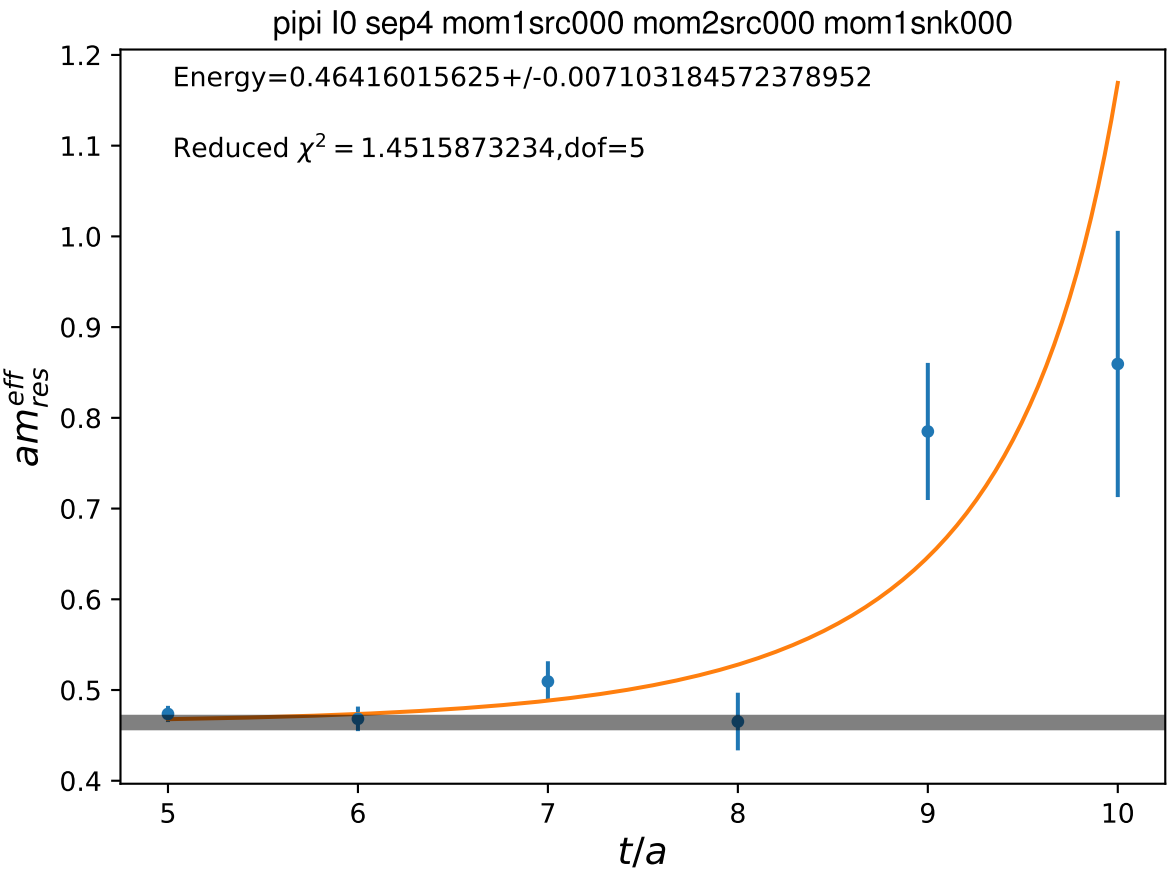

pipi I2 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.510424804687+/-0.0009213997431763147
Reduced $\chi^2 = 0.917112622103$,dof=4
am_{res}^{eff}
t/a
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0

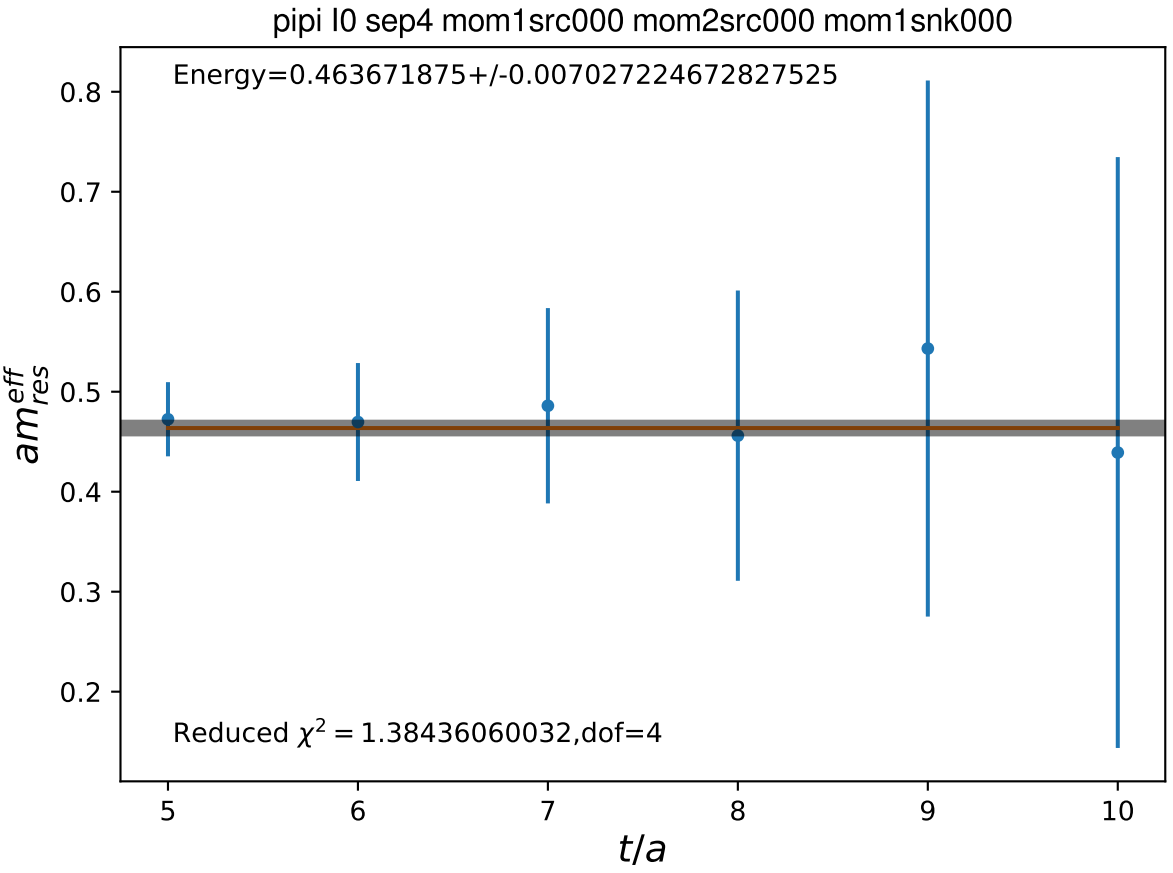

pipi I2 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.510314941406+/-0.0006443437922994932
Reduced $\chi^2 = 0.924087396001$,dof=4
am_{res}^{eff}
t/a
0.518
0.516
0.514
0.512
0.510
0.508
0.506
0.504
6
7
8
9
10
11

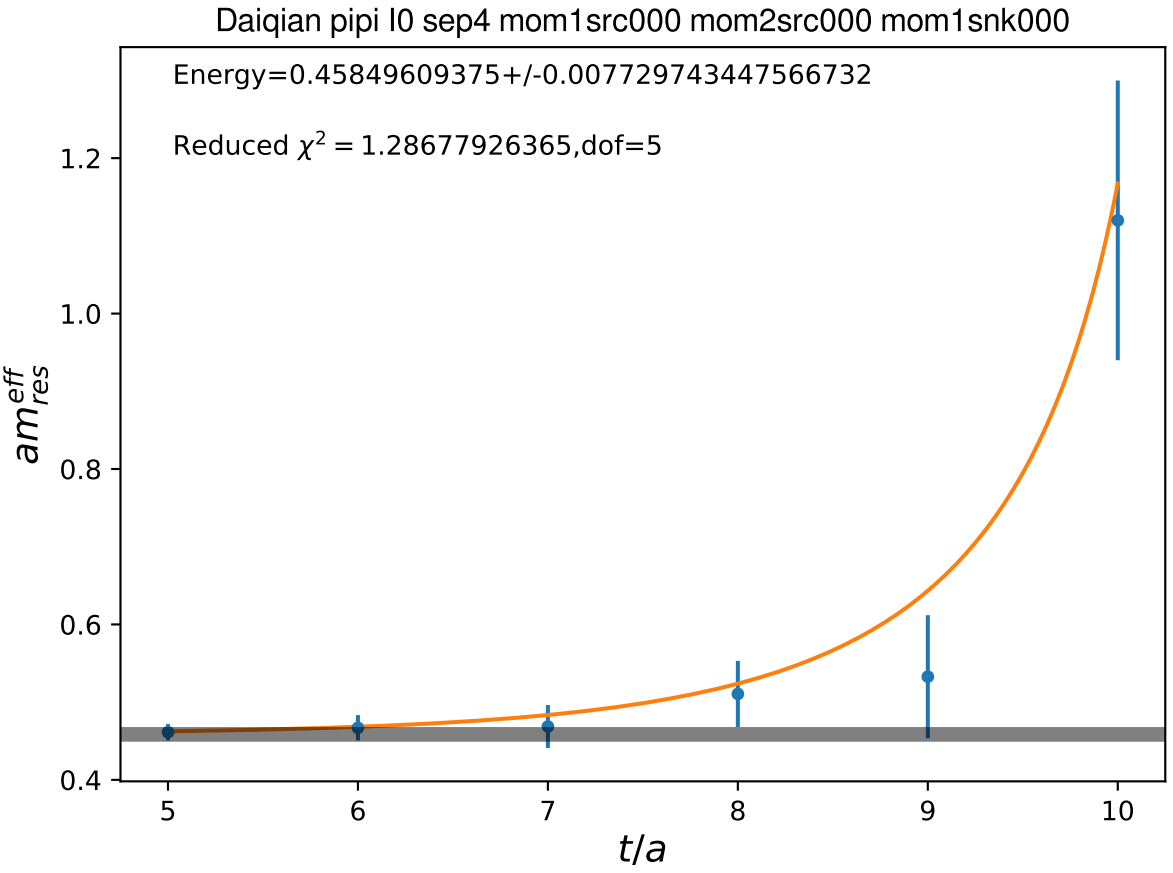

Daiqian pipi I2 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.509155273437+/-0.00125714291205
Reduced $\chi^2 = 0.665707691751$,dof=4
am_{res}^{eff}
t/a
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0

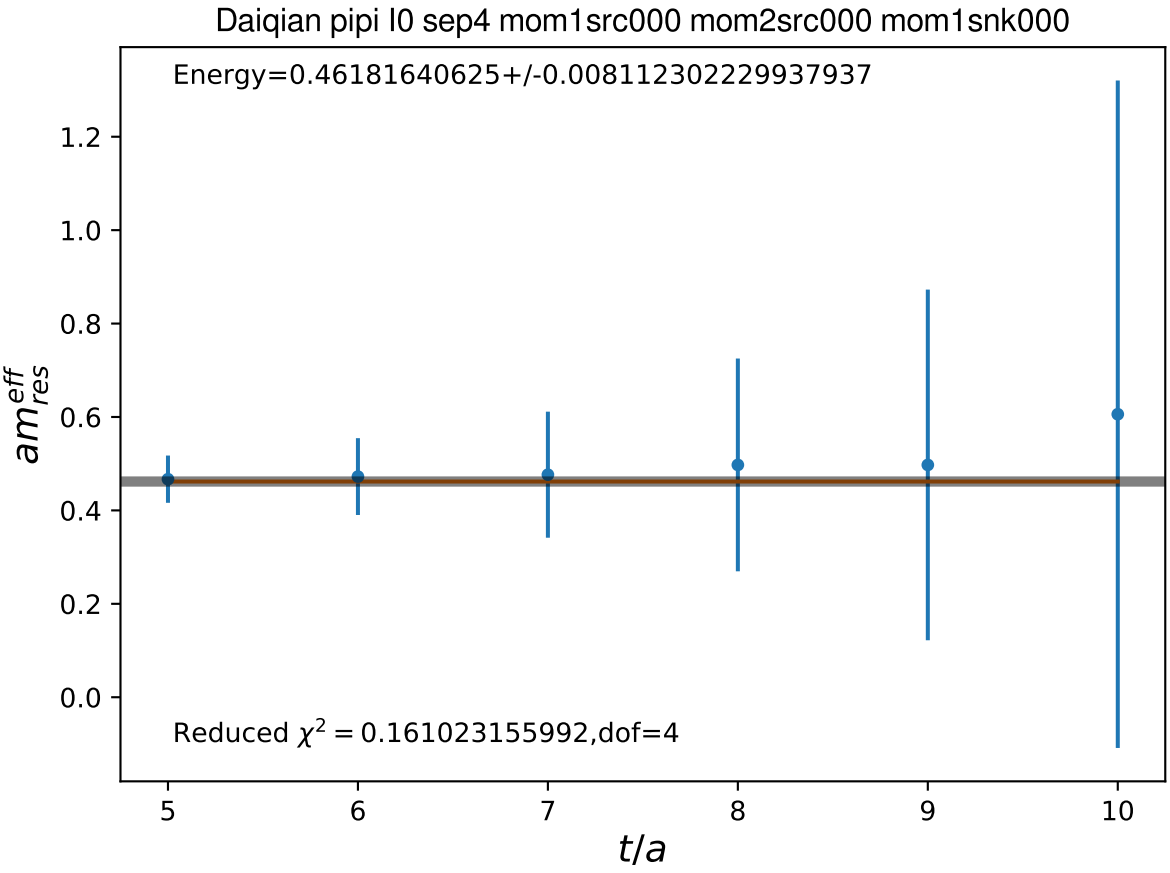

Daiqian pipi I2 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.508959960937+/-0.0008641988791227677
Reduced $\chi^2 = 0.726489293287$,dof=4
am_{res}^{eff}
t/a
0.5200
0.5175
0.5150
0.5125
0.5100
0.5075
0.5050
0.5025
0.5000
6
7
8
9
10
11

B.3.1.3 Pipi, $I = 0$


pipi I0 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.46416015625+/-0.007103184572378952
Reduced $\chi^2 = 1.4515873234$,dof=5
am_{res}^{eff}
t/a
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
5
6
7
8
9
10


pipi I0 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.463671875+/-0.007027224672827525
Reduced $\chi^2 = 1.38436060032$,dof=4
am_{res}^{eff}
t/a
0.8
0.7
0.6
0.5
0.4
0.3
0.2
5
6
7
8
9
10


Daiqian pipi I0 sep4 mom1src000 mom2src000 mom1snk000
Energy=0.45849609375+/-0.007729743447566732
Reduced $\chi^2 = 1.28677926365$,dof=5
am_{res}^{eff}
1.2
1.0
0.8
0.6
0.4
5
6
7
8
9
10
t/a



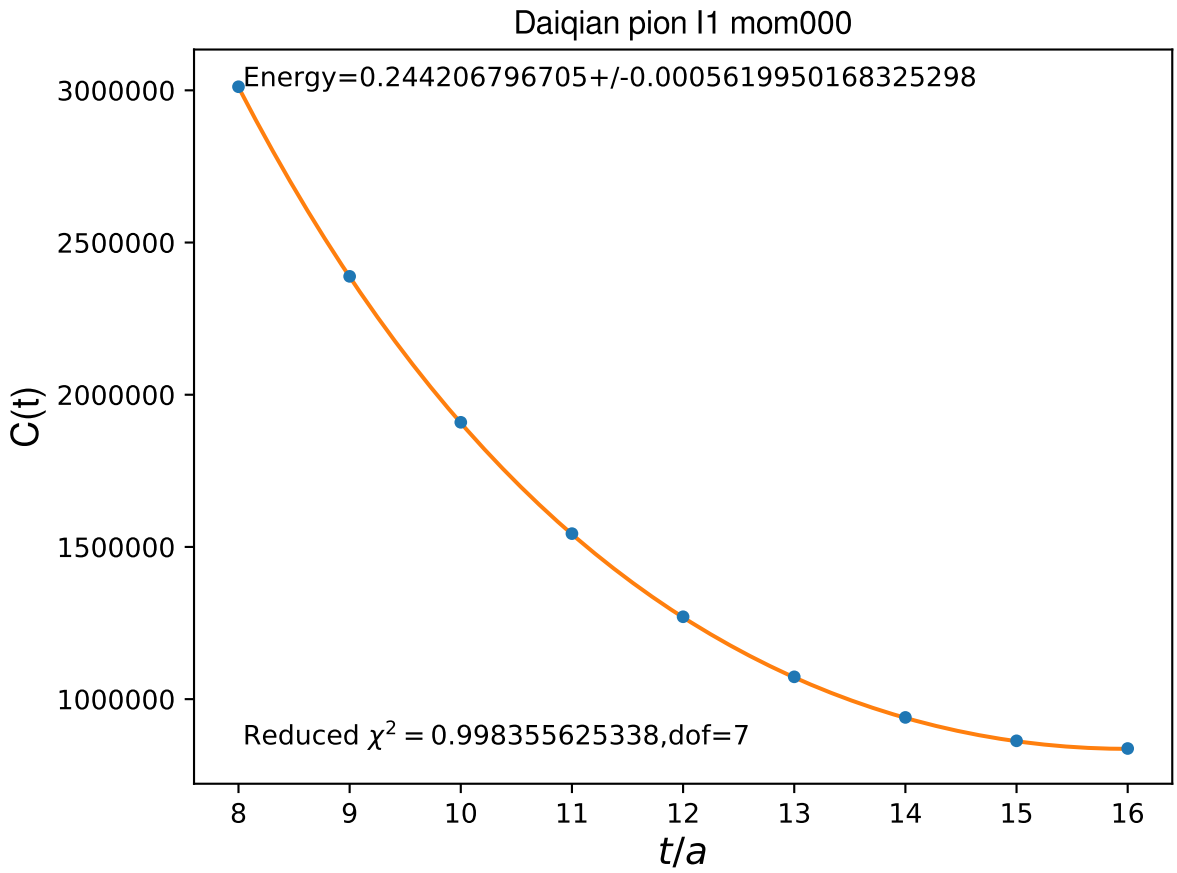
B.3.2 Fits

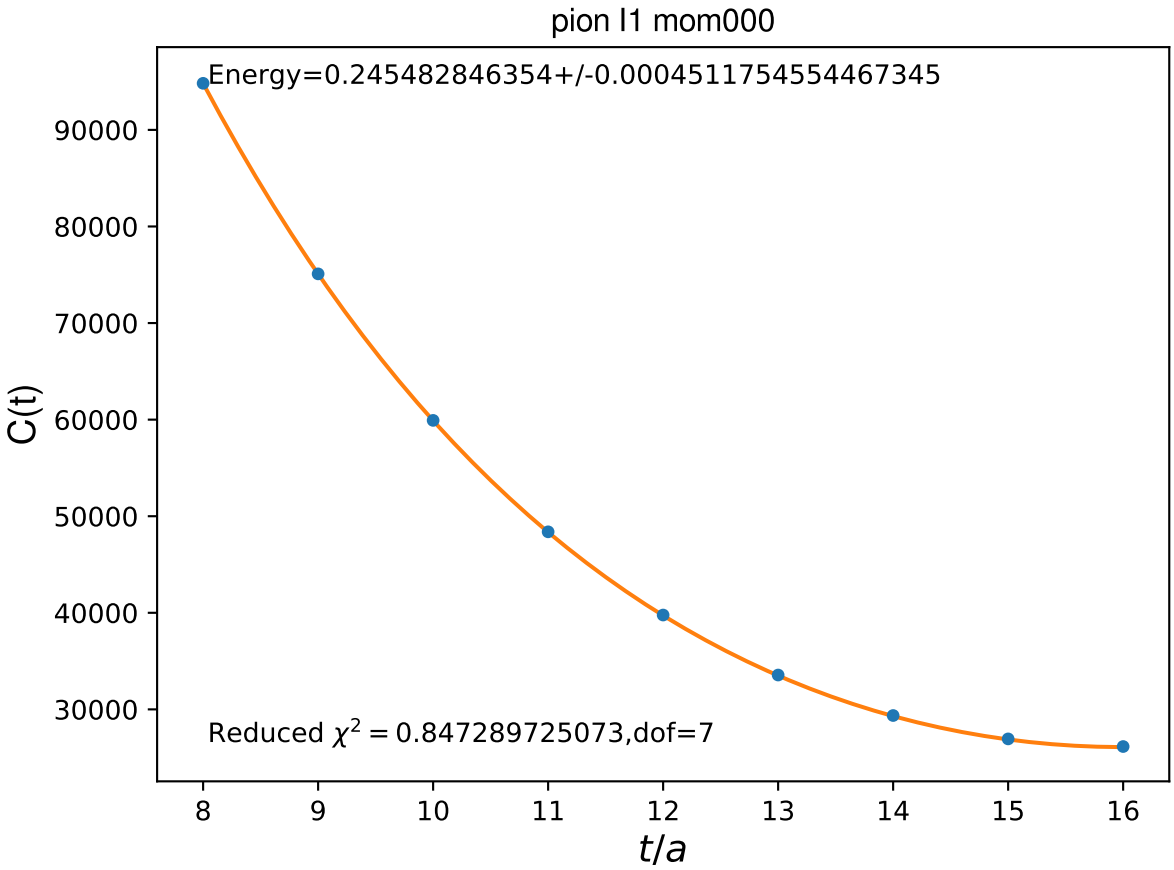
These are single fits are to the average values of the correlation functions. Jackknifing is used to estimate the covariance matrix, but only some of these are jackknife fits. The jackknife fits are annotated as such.

We use Nelder-Mead optimization to minimize χ^2. Up to date implementation details on the fitting routine can be found at [15] (logs of the fits here, including initial guess parameters, are available on request).

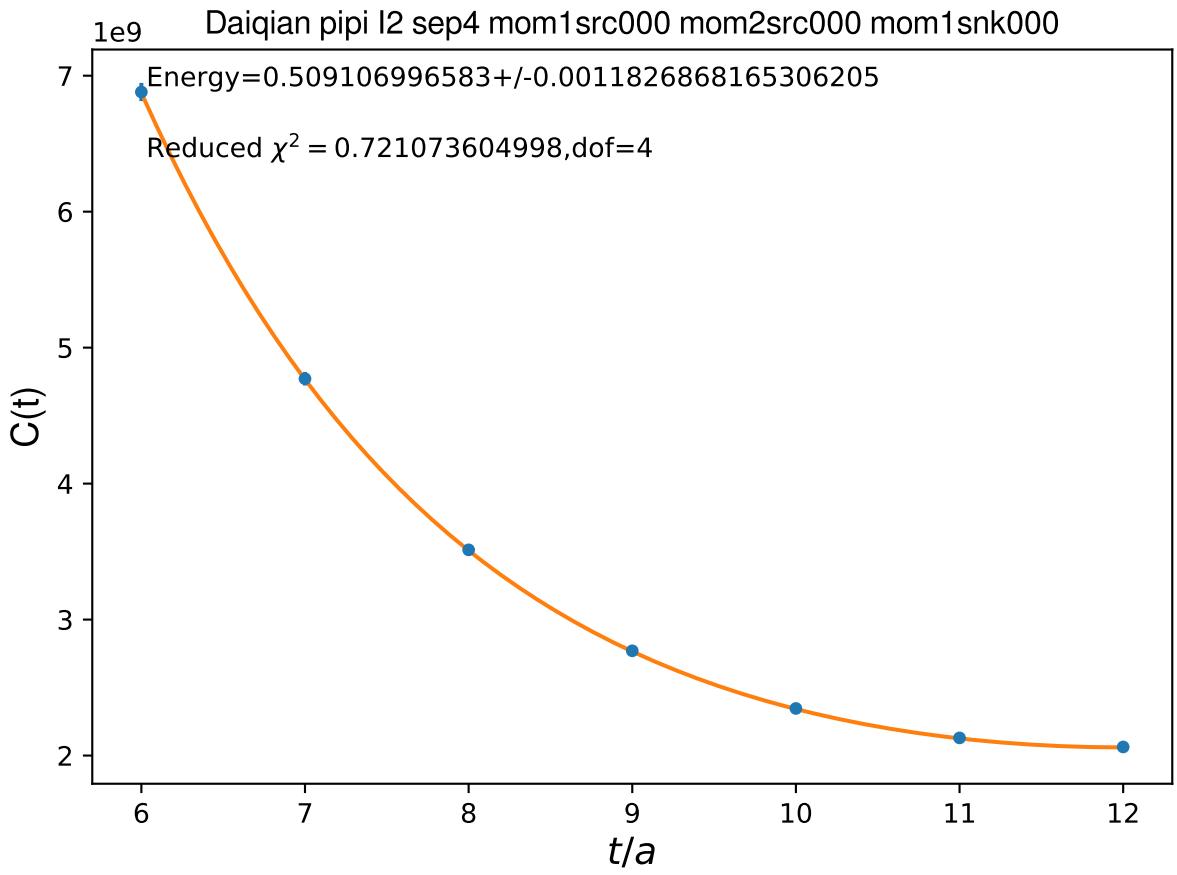
The pion figures are fit to a cosh centered at $t = 16$ whereas the pipi figures are centered at $t = 16$ and have another free parameter in the form of an additive constant.

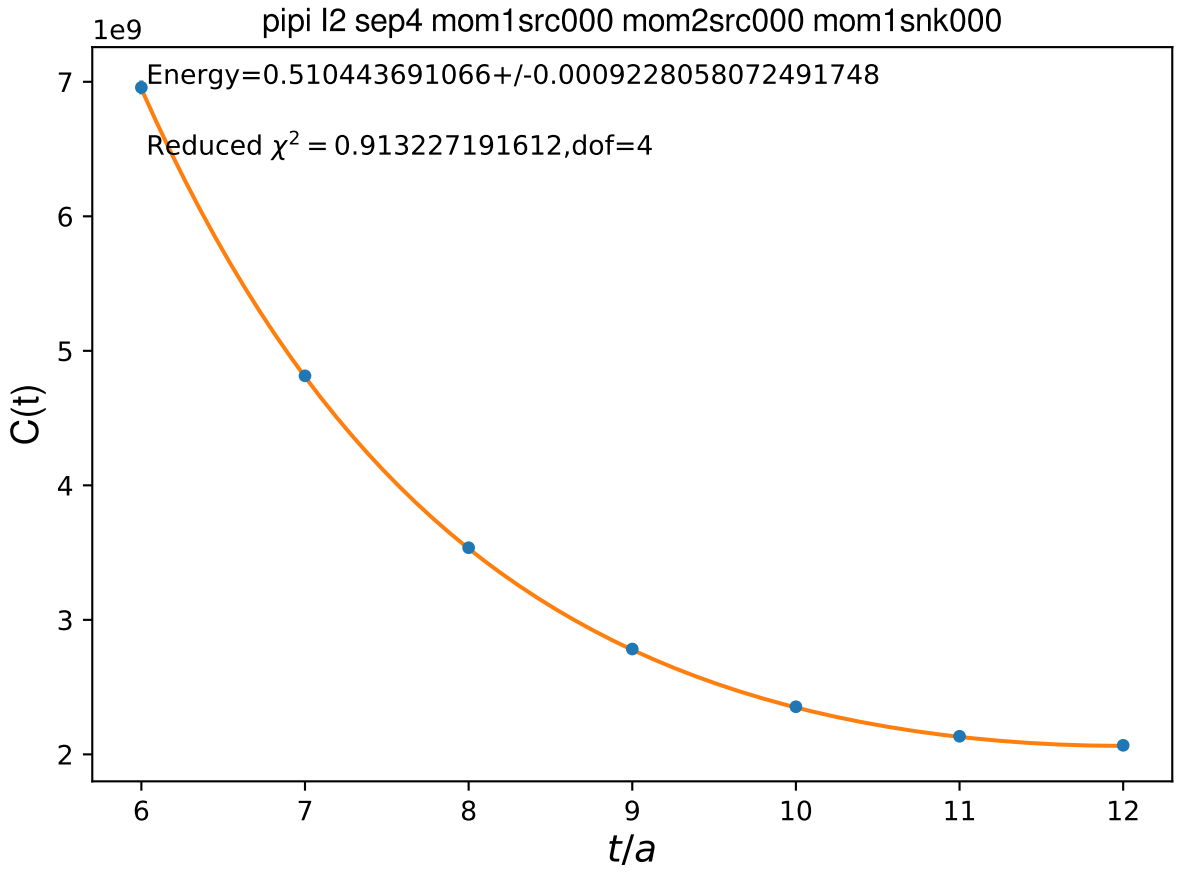
B.3.2.1 Single Pion



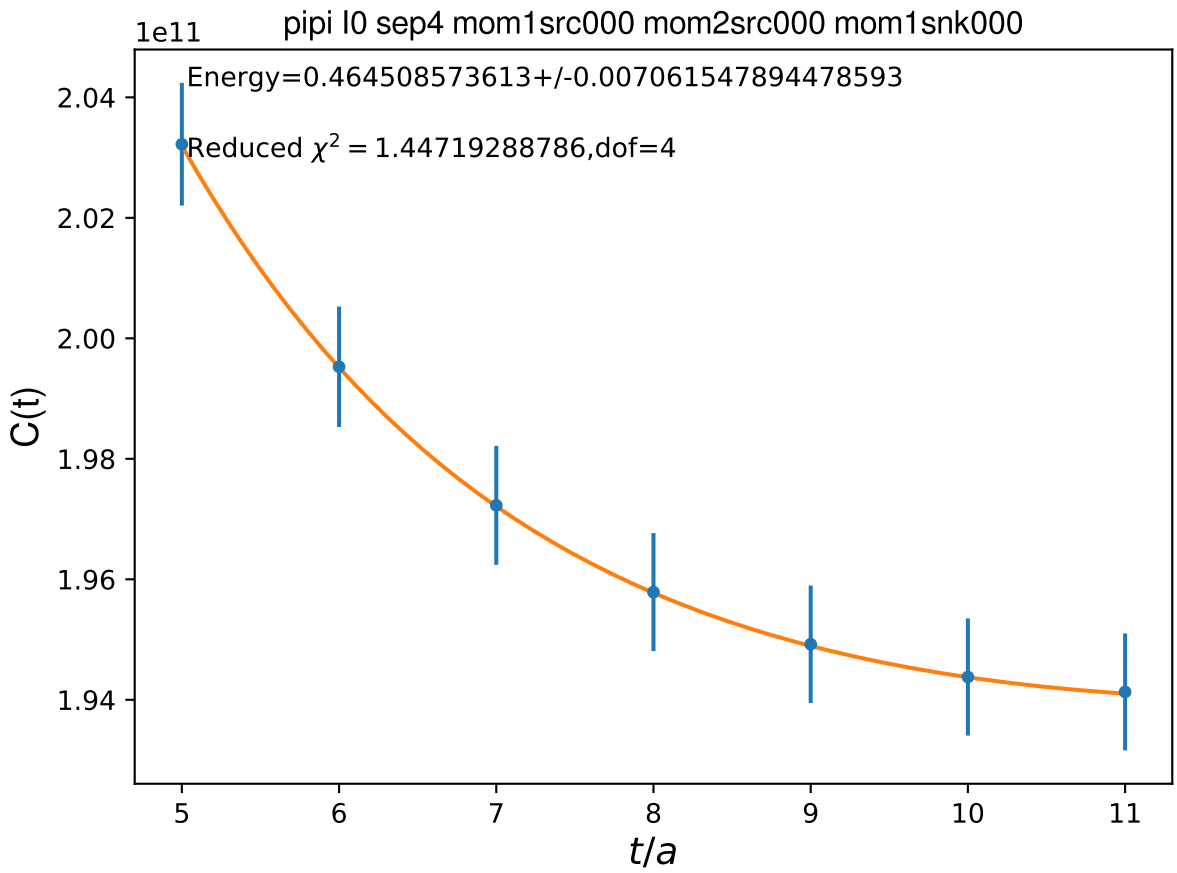

pion I1 mom000
Energy=0.245482846354+/-0.0004511754554467345
Reduced $\chi^2 = 0.847289725073$,dof=7
C(t)
t/a
90000
80000
70000
60000
50000
40000
30000
8
9
10
11
12
13
14
15
16

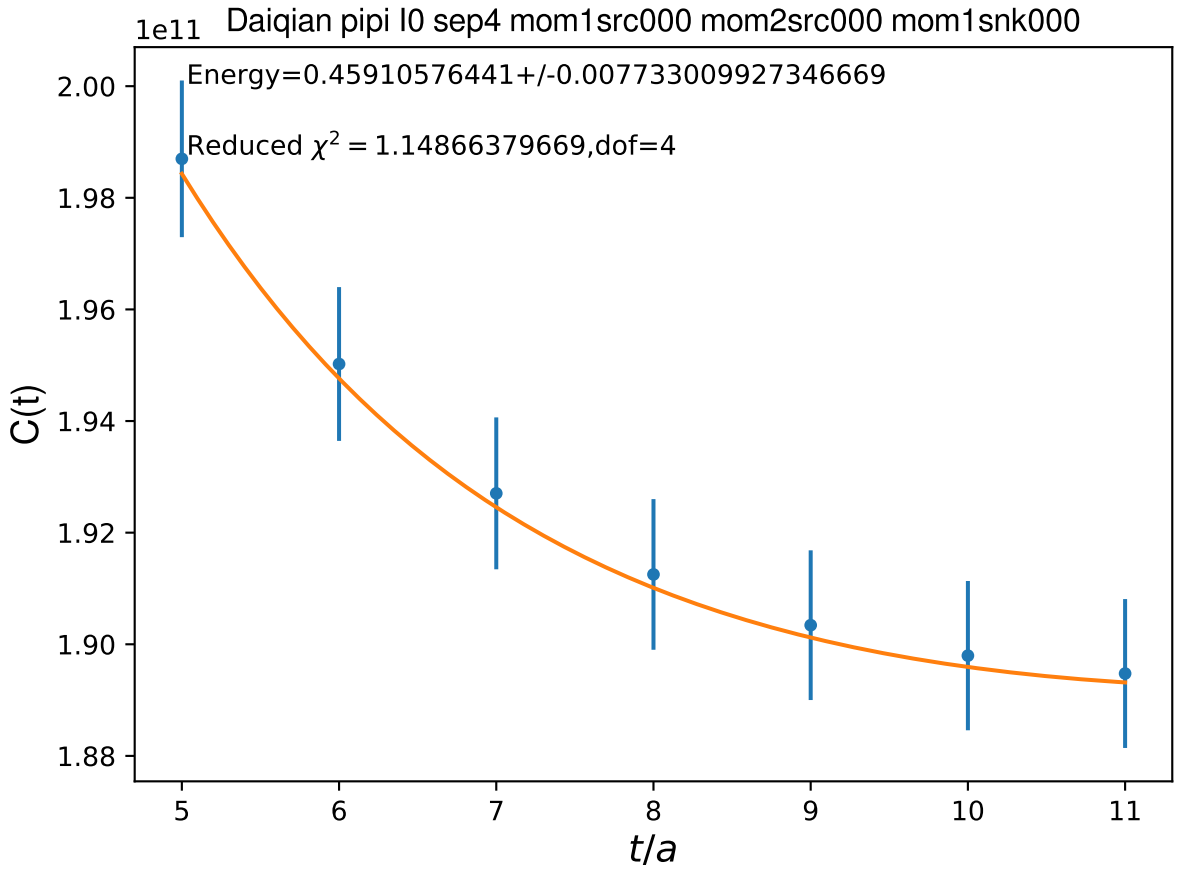
B.3.2.2 Pipi, $I = 2$


Daiqian pipi I2 sep4 mom1src000 mom2src000 mom1snk000
1e9
Energy=0.509106996583+/-0.0011826868165306205
Reduced $\chi^2 = 0.721073604998$,dof=4
C(t)
t/a
7
6
5
4
3
2
6
7
8
9
10
11
12


pipi I2 sep4 mom1src000 mom2src000 mom1snk000
1e9
Energy=0.510443691066+/-0.0009228058072491748
Reduced $\chi^2 = 0.913227191612$,dof=4
C(t)
t/a
7
6
5
4
3
2
6
7
8
9
10
11
12

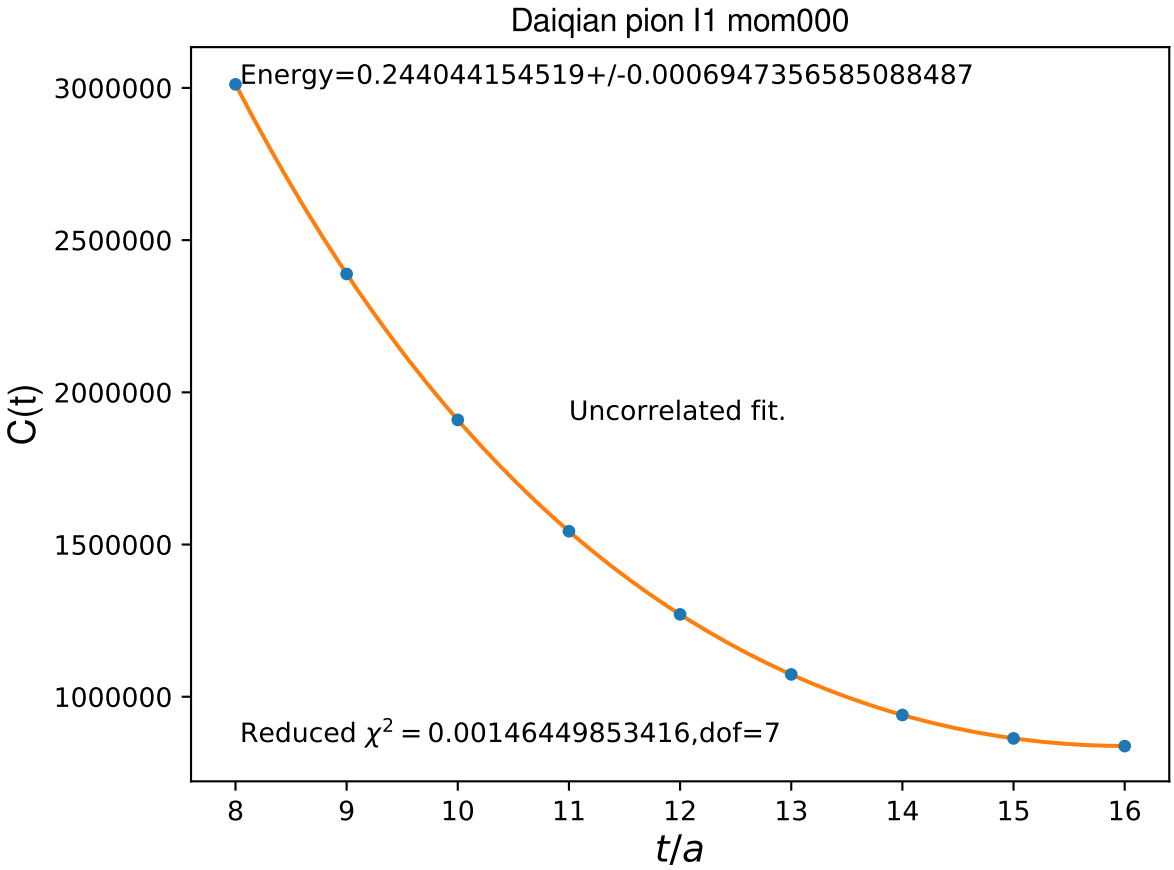
B.3.2.3 Pipi, $I = 0$

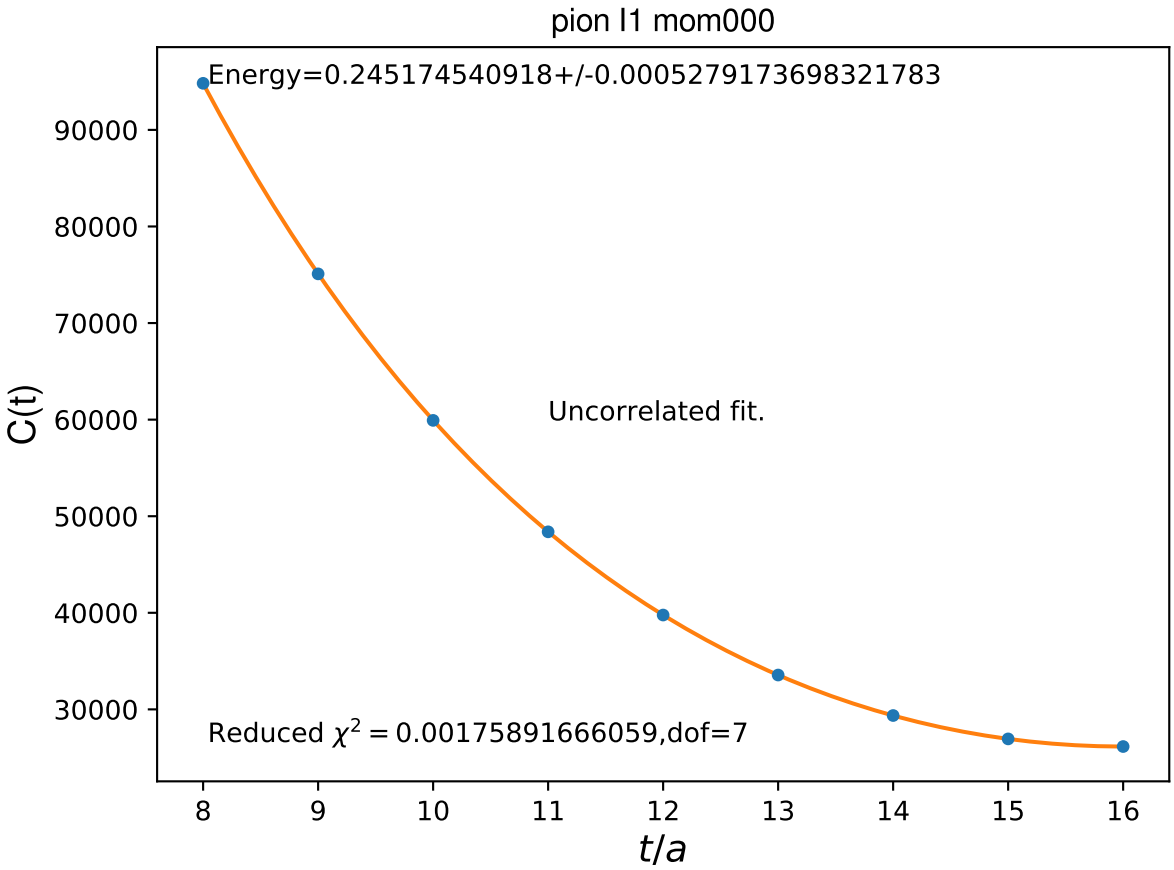

pipi I0 sep4 mom1src000 mom2src000 mom1snk000
1e11
Energy=0.464508573613+/-0.007061547894478593
Reduced χ^2 = 1.44719288786,dof=4
2.04
2.02
2.00
1.98
1.96
1.94
C(t)
5
6
7
8
9
10
11
t/a

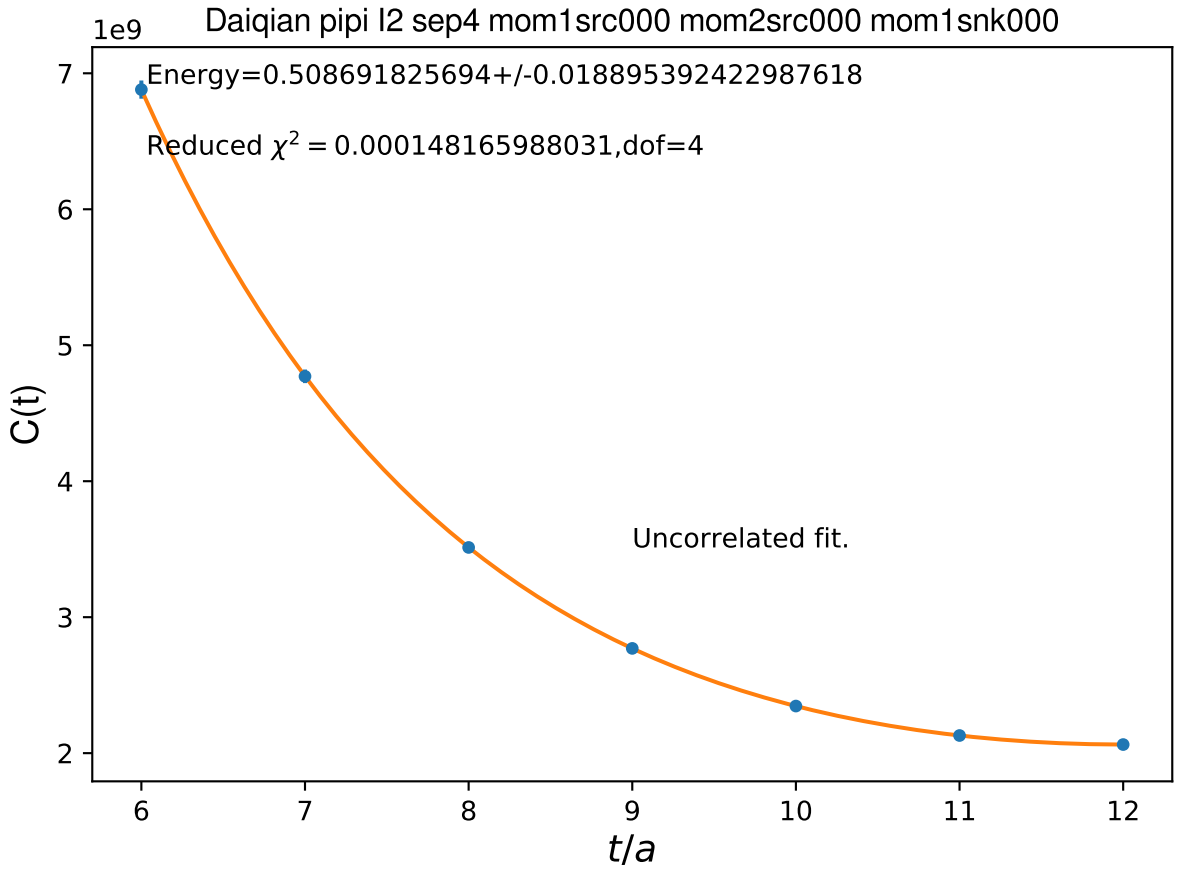


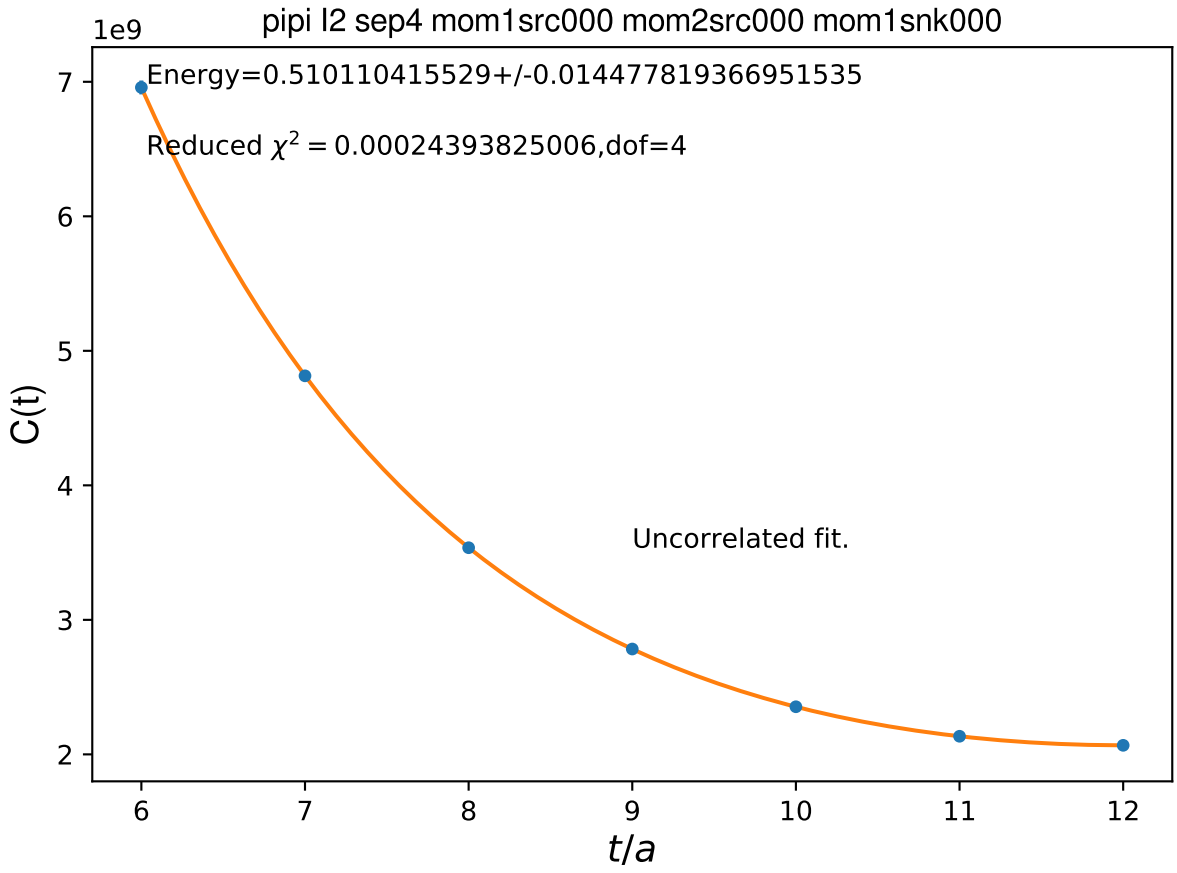
B.3.2.4 Uncorrelated Fits to C(t)

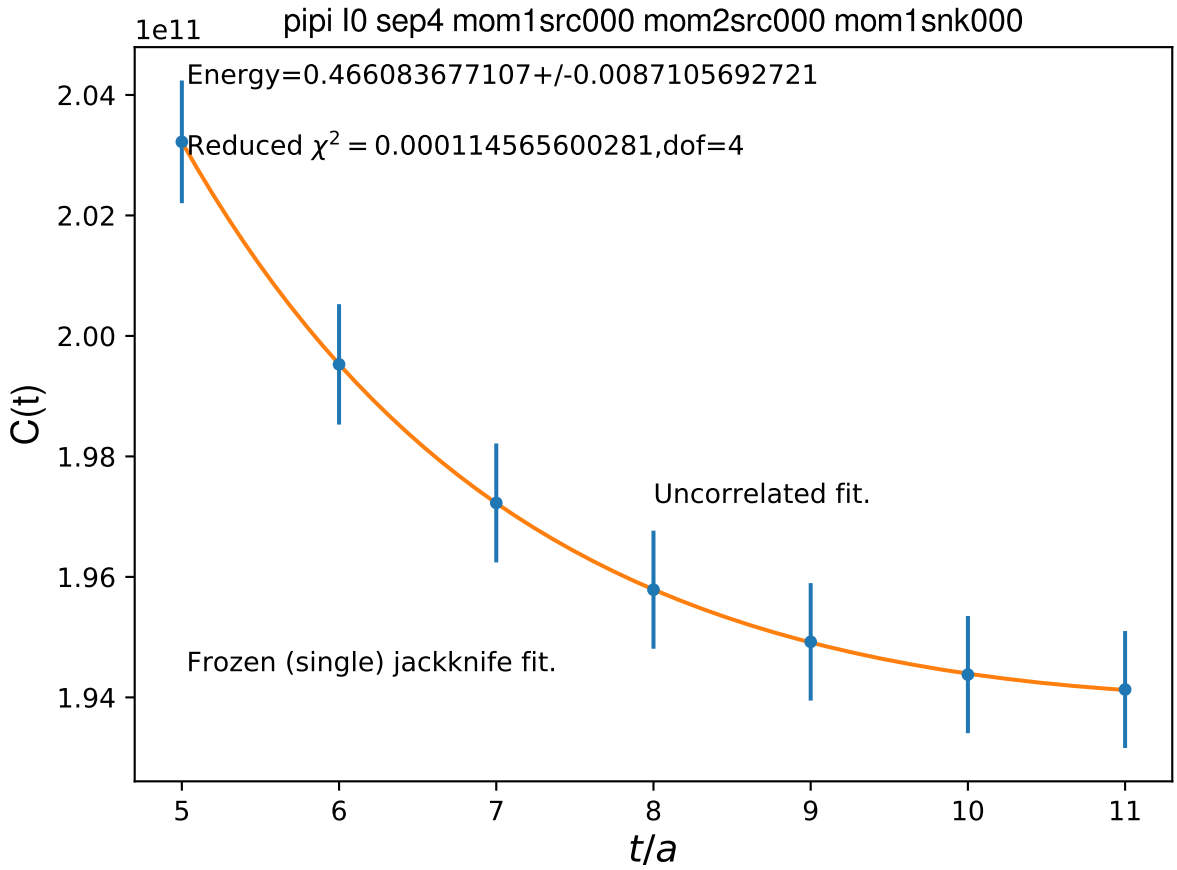
Unless otherwise noted on the diagram itself, fits are correlated.

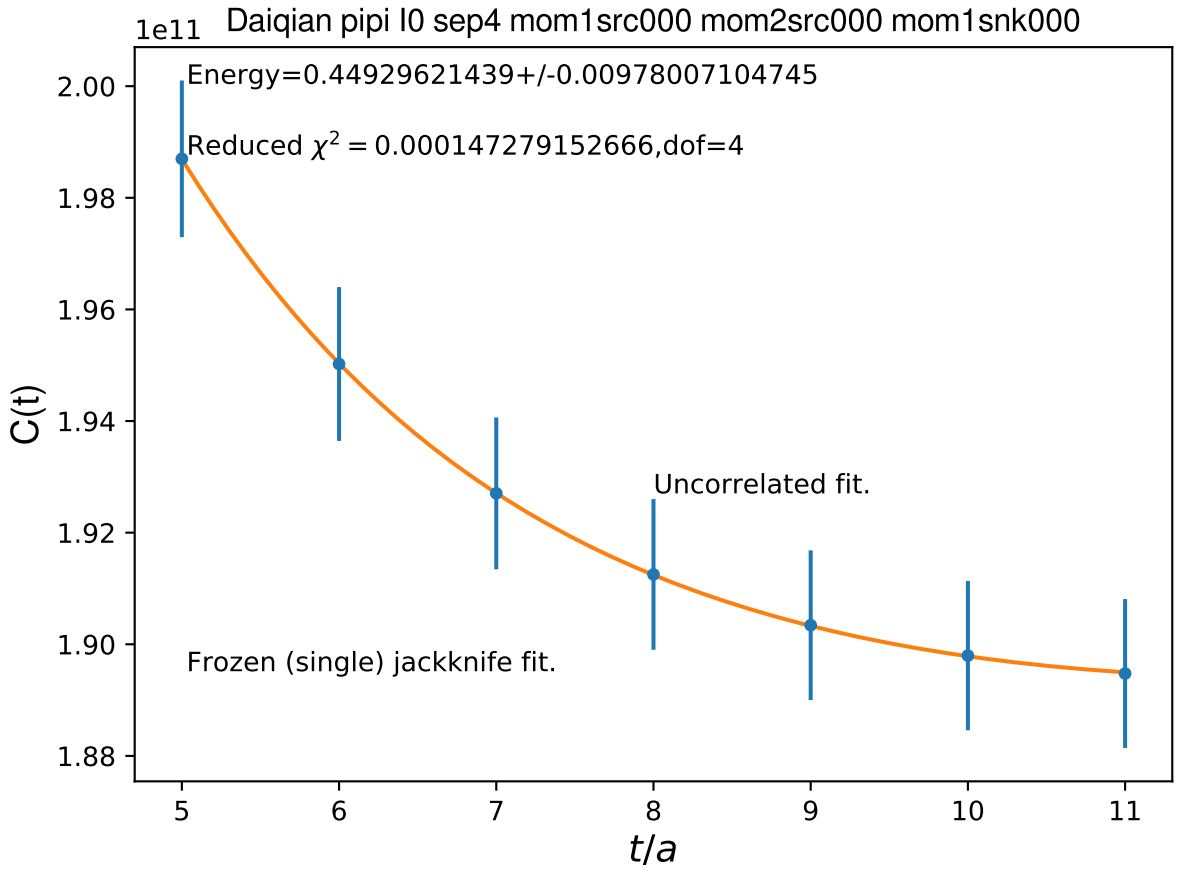

Daiqian pion I1 mom000
Energy=0.244044154519+/-0.0006947356585088487
Uncorrelated fit.
Reduced $\chi^2 = 0.00146449853416$,dof=7
C(t)
3000000
2500000
2000000
1500000
1000000
8
9
10
11
12
13
14
15
16
t/a


pion I1 mom000
Energy=0.245174540918+/-0.0005279173698321783
Uncorrelated fit.
Reduced $\chi^2 = 0.00175891666059$,dof=7
C(t)
t/a
90000
80000
70000
60000
50000
40000
30000
8
9
10
11
12
13
14
15
16


Daiqian pipi I2 sep4 mom1src000 mom2src000 mom1snk000
1e9
Energy=0.508691825694+/-0.018895392422987618
Reduced χ^2 = 0.000148165988031,dof=4
Uncorrelated fit.
C(t)
t/a
7
6
5
4
3
2
6
7
8
9
10
11
12


pipi I2 sep4 mom1src000 mom2src000 mom1snk000
1e9
Energy=0.510110415529+/-0.014477819366951535
Reduced χ^2 = 0.00024393825006,dof=4
Uncorrelated fit.
C(t)
t/a
7
6
5
4
3
2
6
7
8
9
10
11
12


pipi I0 sep4 mom1src000 mom2src000 mom1snk000
1e11
Energy=0.466083677107+/-0.0087105692721
Reduced $\chi^2 = 0.000114565600281$,dof=4
Uncorrelated fit.
Frozen (single) jackknife fit.
C(t)
t/a
2.04
2.02
2.00
1.98
1.96
1.94
5
6
7
8
9
10
11



B.4 Conclusion

One could expand on this analysis by looking at binned diagrams, or fitting under a jackknife (double/single jackknife, instead of single fit to average).

Some good agreement is obtained between Daiqian's early A2A run and the more up to date calculation done on the new KNL. Better agreement would be possible only if the random number seed/method were known.

The errors in the mass for our data may be artificially smaller because the configurations are not spaced out very far in monte-carlo time (mostly around 5). However, binning with sizes of $20, 40, 60, 80$ has only a negligible effect (for correlated fits) on our data and Daiqian's data (although need to check with corrected fitter). Binning does still needs to be checked again, though.

Versions of this chapter can (at the very least could) be tracked via the local git repository under the individual postings on the RBC Columbia server. If you do not have access to this server and would like a more complete document history, feel free to e-mail the author (daniel.hoying@uconn.edu) for a copy.

Appendix C

Data Format Guide

C.1 Location

Assuming one has access to JLab's computers, the data is located under /cache/K2pipiPBC/qcddata. The 24^3 data is located under a folder called new_pipi. The original "pipi" folder has spin contamination because the overall irrep projection is missing correlation functions (bug since fixed). There is extra data located under fill_in folders. These data are from runs which did not finish under the wall clock limit. Exact configs (for section 2.6) are located under the "exact" folder.

The 32^3 data is similar, but instead of a fill_in folder, there are "partial" folders. Thus, if a config is under partial, it is also under the main folder (and both datasets should be combined to obtain the complete set of correlation functions). "partial_fixed" in the main folder should be preferred as it has corrected $t_{\text{dis,max}}, t_{\text{step}}$. The "exact" folder, of course, has the exact configs.

C.2 Compiling the Data

Under each config folder, there is a "props/output" folder which has all of the correlation functions stored in hdf5[45] files. These are labeled by the config number and an index. This index is not meaningful except to make the files unique. The idea was to have a fixed number of correlation functions per hdf5 file to avoid data loss should the job exit early due to something like a node failure. There are files also labeled as "discon." These have the bubbles (which contain a single time index).

One should form an hdf5 file of links linking all of these files together into an overall hdf5 file for one config. The datasets themselves can be converted into numpy[43] arrays as soon as the file is opened with h5py[45].

C.3 Dataset Naming Conventions

The dataset naming conventions should map fairly clearly into diagram classes defined in [7]: (C,D,R). A label of vec indicates $I = 1$ contractions while a label of $scalar$ indicates σ contractions. There is label of sep which corresponds to t_{sep}. t_{sep} is the distance between the inner and outer pions in time.

C.3.1 $\pi\pi \to \pi\pi$; General Conventions

For $\pi\pi \to \pi\pi$, the momentum combinations are labelled $mom1src <$ inner source pion momentum $> \_mom2src <$ outer source pion momentum $> \_mom1snk <$ inner sink pion (physical) momentum $>$ The momentum strings have underscores to represent minus signs. Thus, for example, _101 or 0_11, give, respectively, $(-1,0,1)$ and $(0,-1,1)$. The sink momentum in the file name is the physical momentum, so a file name which ends in, e.g.,

$$mom1src001\_mom2src010\_mom1snk010$$

would describe the physical process $\pi(0,0,1)+\pi(0,1,0) \to \pi(0,1,0)+\pi(0,0,1)$.

C.3.2 Single Particle Correlation Functions

$Hbub$ diagrams are for single particle correlation functions. The two types are the $\rho \to \rho$ correlation functions and (connected) $\sigma \to \sigma$ labeled by vec and $scalar$ respectively. These have a single momentum which labels the physical momentum of the correlation function. The polarization info is $src-snk\_<\mu>\_<\nu>$ where μ,ν correspond to the bilinears $\overline{\psi}\gamma_\mu\psi$,$\overline{\psi}\gamma_\nu\psi$.

There are also single pion correlation functions which are 'pioncorr' and 'pioncorrChk'. 'pioncorr' has the proper normalization but skips source time slices according to t_{step} (see section C.4.1). 'pioncorrChk' (the name is due to historical accident) has $t_{\text{step}} = 1$ (all source times), but has the wrong normalization. Multiply the correlation function by 2 to get the proper norm (as one can see comparing the time slices in common between 'pioncorr' and 'pioncorrChk').

C.3.3 $\rho,\sigma \to \pi\pi$

T figures are for the processes $\rho \to \pi\pi, \pi\pi \to \rho, \sigma \to \pi\pi, \pi\pi \to \sigma$. These have two momenta: one source and one sink. The source momenta labels either the momentum of the ρ,σ or the inner source pion. The sink momentum labels the physical momentum of either the inner sink pion or the ρ,σ. The R which appears in the σ T correlation functions implies that this is a diagram with σ at the source location. Thus, $FigureT\_scalar\_momsrc010\_momsnk000$ corresponds to $\sigma(0,1,0) \to \pi(0,0,0)+\pi(0,1,0)$. The $\rho \to \pi\pi$ vs.

$\pi\pi \rightarrow \rho$ correlation functions are distinguished by the location of the ρ's polarization label (whether source (src) or sink (snk)).

C.3.4 Disconnected Diagrams

The disconnected diagrams are located in separate hdf5 files (as mentioned) and have labels like $Vdis$ or $scalar-bubble$ indicating a $\pi\pi$ bubble and σ bubble respectively. We need to multiply these together to form the disconnected diagrams which enter the into the isospin projections (see sections 4.3.1 and 4.3.3).

C.3.5 T: $vecCheck$ Diagrams

We know the two topologies of the ρ vector T must come with a relative minus sign, but we also saved the (incorrect) version with a relative plus sign under the label $vecCheck$ in case we wanted to access the individual topologies of T.

C.4 Array Shape vs. Time

The connected diagrams have shape (L_t,L_t), which correspond to $t_{\text{src}}, t_{\text{dis}}$. This data is (clearly) un-averaged over t_{src} and unfolded. t_{dis} is the distance between inner pions or inner pion and ρ, σ.

The disconnected diagrams have shape (L_t). corresponding to the either the location of the σ bubble or the location of the later pion in time. The other pion in the $\pi\pi$ bubble is located at $t - t_{\text{sep}} \% L_t$ (see section C.3 for definition of t_{sep}).

C.4.1 t_{step}

We skip source times for the connected diagrams to save on compute time. We know these extra times do not help our statistics much for these diagrams (perhaps due to correlations). In the case where t_{step} divides L_t, we use all source time slices such that $t_{\text{src}} \% t_{\text{step}} = 0$ (e.g., section 3.2). Otherwise (as in section 3.3), we use $t_{\text{src}} \% t_{\text{step}} = 0$ except if the maximum t_{src} subtracted from L_t is less than t_{step}. Thus, for 32^3 and $t_{\text{step}} = 10$, we use $t_{\text{src}} = \{0, 10, 20, 30, 40, 50\}$.

C.5 Old Data

(run on ensemble detailed in section 3.1)

C.5.1 Location

The data is available at JLab under /cache/K2pipiPBC/qcddata/DWF/2+1f/16nt32/run*.

C.5.2 Intro

This document serves to explain the data format of old Brookhaven KNL data in case the reader would like to do their own analysis. It also describes some caveats and where to find the data.

C.5.3 Description; Caveats

Metadata is located in the numbered folders under run[1-7]. These include vmls, run scripts, and slurm output. The data itself is located under run[1-7]/collect in tarballs (except for run6 which has some of these gzip'd). The data in runs 1-5 has (maximum) one unit of momentum (this limit applies to center of mass and relative momenta). run6 has 2 units, but 0 time seperation (everything else has time sep. 4).

C.5.4 Bugs, Caveats

In the course of run7, bugs were discovered in how the momenta were calculated. One bug is that the momenta at sinks are the negative of the labeled sink momenta. This means all diagrams with non-zero center of mass below (and here the meaning of below will need to be clarified) a certain config in run7 do not conserve momenta. This bug hits all diagrams, except for the pioncorrChk diagrams, which should be good for all runs and momenta.

The second bug was a typo in the momenta for T diagrams. The result of this typo was that one of the pions was fixed to have the same momenta as the ρ/σ at sink. Coupled with the above bug, this means all T diagrams below the boundary config with any momenta do not conserve momenta.

The boundary config is 6260 (which is bad). Anything including this config and below have these bugs. Importantly, because the configs ran in "alphabetic" order, below does not mean integers <=6260. The good configs are, instead, any configs in which the first two digits are >= 63 (and 6280 is good). This gives around 140 configs in which all momenta are conserved.

This does not mean that the data below this config is uniformly bad, but it does mean that one should swap the sink momentum label for $\langle\pi\pi|\pi\pi\rangle$ diagrams, avoid non-zero center of mass, and avoid T diagrams with momenta. This still gives quite a few analyses that might be done.

C.5.5 Format

Each diagram is a separate ASCII file (we later switched to hdf5 as to handle the large number of momentum combinations which generated an unwieldy number of small files). If the diagram is not a disconnected bubble, there should be 32^2 lines per file. The format is

$$tsrc\ tdis\ Re(Corr)\ Im(Corr)$$

tsrc is the time of the inner source particle. tdis is the distance in time between inner sink and inner source. Inner particles consist of one particle at source and one at sink which have the minimum time separation (if we ignore wrap-around). Thus, the inner particle at source is at the later time of the two source particles and inner particle at sink is the earlier of the sinks.

If the file is a disconnected bubble, there are 32 lines. The bubble is a trace calculated on a particular time slice. These are then multiplied at the analysis stage to get the full disconnected diagram (and allows vacuum subtraction). The format here is

$$t\ Re(Corr)\ Im(Corr)$$

Each diagram is described fully by its filename. The beginning of each file is

$$traj\_ < trajectory\ number > \_[...]$$

followed by a string for the figure name. If the string is like Figure_, this means the file is disconnected bubble.

The string sep<number> indicates the time separation between two source pions (also the time separation between two sink pions).

C.5.5.1 $\langle \pi\pi | \pi\pi \rangle$ Diagrams

Diagram Names:

FigureR,C,D are the typical connected diagrams. In particular, they are 0.5 times the quantities in parentheses of section 3.3.1 of chapter 4 (and 0.5 times the diagram on the far right of this sum). These are consistent with the conventions established by Qi Liu ([7]) for C, D, and R.

This format is so far completely consistent with C. Kelly's format (private comm.).

Figure_Vdis is the $\pi\pi$ bubble. The time in this file is the time of the later pion. However, each bubble has a 0.5 multiplying it, so one needs to multiply all of these values by a correction factor of 2. This was later fixed so the new norm is 1.

If FigureR or FigureD has a 'vec' in its filename, then we are actually referring to the versions of FigureR and figureD in $I = 1$. FigureR vec is -0.5 times the quantity in parentheses of section 5 in chapter 4. FigureD vec is 0.5 times the difference of the two diagrams outside of the parentheses (i.e. they have a relative minus sign).

C.5.5.2 $\langle \rho|\pi\pi \rangle$ Diagrams

There is one (relevant) diagram here: FigureT_vec. As in the $\pi\pi$ case, the vec indicates $I = 1$, and a relative minus sign between the two topologies (see chapter 5. This diagram is 0.5 times the amplitude in chapter 5 for $\langle \rho|\pi\pi \rangle$). The diagrams that are vecCheck is the same amplitude with the relative minus sign between the two topologies replaced by a plus sign.

The string pol_snk_<number> indicates the polarization of the sink particle (the ρ, in this case.) The polarization is either $1,2$ or 3 corresponding to, respectively, x, y, z.

momsrc is the inner (source) pion momentum. momsnk is the momentum of the ρ. The momentum of the outer source pion is given implicitly by momentum conservation.

The reverse amplitude, $\langle \pi\pi|\rho \rangle$ is an auxiliary diagram (see section 2.1).

C.5.5.3 $\langle \rho|\rho \rangle, \langle \sigma|\sigma \rangle$ Diagrams

These diagrams are in FigureHbub diagrams. The momentum is the center of mass momentum (hence is the momentum at the source and sink). The σ diagrams are in FigureHub_scalar.

The vec diagrams are ρ's. The polarization info is pol_src-snk_<number1>-<number2>. number1 is the polarization at the source. number2 is the polarization at the sink.

The diagram is 0.5 times the $\langle \rho|\rho \rangle$ amplitude (same for σ) (hence one needs a correction factor of 2). This was later fixed, so the newer data has norm 1.

C.5.5.4 $\langle \sigma|\pi\pi \rangle$ Diagrams

One of the diagrams is a fully connected FigureT labeled with 'scalar' instead of 'vec' for the ρ. Again, this diagram is the average of the two topologies just like in the ρ case (the two topologies are the same up to particle flow direction). Thus, this diagram is 0.5 times the triangular diagram given in chapter 4.

The other diagram is disconnected. One half is made from the usual $\pi\pi$ bubble. The other half is in the file 'Figure_scalar-bubble' with momentum given by mom<usual momentum string>. There is no correction factor

needed for this bubble.

C.5.5.5 Pion correlation function

The single pion two point should be gotten from $pioncorrChk$ diagrams (the other non-Chk diagrams are identical except for diagrams that aren't mom000 below the run7 cutoff 6260. These do not conserve momentum). The momentum is given by mom<usual mom. str.>

The amplitude given is 0.5 times $\langle\pi|\pi\rangle$ (for both Chk and non-Chk).

C.5.6 Conclusion; Aux diagram note

Please note that in these early runs auxiliary symmetry was nominally used (see section 2.1). However, it was subsequently discovered in the early 24^3 (section 3.2) runs that significant numbers of momentum combinations were missing due to a bug in the implementation of aux symmetry. Thus, irrep projections will necessarily be incomplete on this data set.

Appendix D

Building the Production Toolset

Assuming one wants to run a reproduction study of any of the $\pi\pi$ data, this section covers the compilation procedure on the KNL.

Located under cps_pp, one runs cfig.sh (laptop builds), cfig_bnl.sh (KNL at Brookhaven Lab builds), or cfig_jlab.sh (KNL at JLab) to configure. One then can run make to compile the library. Under cps_pp/work/pion2pt, run make again to compile the binary.

Now, we look at compiler/Grid options used. Under the main directory, create a build directory, then run bootstrap.sh. If one is using our Grid version [53], one might have to uncomment the wget command in bootstrap.sh (or obtain the eigen source some other way). One then calls make[28] to make the library.

Table D.1: Grid Compiler Options

Configure Options
../configure –enable-precision=double –enable-simd=KNL –enable-comms=mpi-auto –enable-mkl CXX=icpc MPICXX=mpiicpc –prefix=/u/home/dsh/knl/install_mpi CXXFLAGS='-DMAP_HUGETLB -DMADV_HUGEPAGE' LIBS='-lrt'

Appendix E

Performing the Analysis (Misc. Notes)

The following are some miscellaneous notes on the analysis. These steps should not be considered exhaustive.

E.1 Fit Procedures

1. We start with production hdf5 files which are labeled: traj_<trajectory number>_<N>.hdf5. $K->\pi\pi$ files and disconnected π,σ bubbles get an extra label, e.g. traj_1000_0_discon.hdf5.

2. We then use a feature of the hdf5 python library h5py which allows us to link all the files in one meta-file of links which can be accessed in the same way as the individual hdf5 files, but is very small and has all the hdf5 datasets for an individual trajectory. See h5link.py in [54] (program we use)

3. We then process the datasets into single elimination jackknife blocks, average over t_{src}, fold in time ($f(t) \rightarrow \frac{1}{2}\left(f(t)+f(T-t-2*t_{\text{sep}})\right)$) ($T$ is the length of the lattice in time; see section C.3 for definition of t_{sep}), and project them onto their respective operators. This procedure currently takes about (e.g.) 2000 s on my laptop running in parallel on 4 intel i7 cores for 155 configs. When we start utilizing aux symmetry this number will jump up by a lot due to an existing redundancy where the hdf5 files are read twice. This is an open problem, but is straightforward to solve. Now we have operators which are single dataset hdf5 files with shape (<number of trajectories>, <Lt>). These are labeled by <operator i><operator j>_<irrep>.jkdat with i,j the matrix indices of the GEVP. See op_compose.py in [54] for the operators we use.

4. Each distinct topology (way of doing a contraction) is averaged over for a given diagram. To fix this norm, multiply all figure 'R' diagrams by 4 (since there are 4 topologies: 2 different arrow directions

times 2 different sink pions the inner pion can contract with). Figures 'D' and 'C' need a factor of 2. See section 4.6 for further details.

5. We choose the fit window and plot window by specifying as command line options –xmax, –xmin, –fitmin, –fitmax and an integer time slice. The error handling is fairly robust on these options, and the fitter knows not to fit outside the plot range. It automatically reduces the fit range in a reasonable way if there is a command line mismatch.

6. We read in the data into the fitter and perform our selected ATW subtraction.

7. We solve the GEVP, enforcing hermiticity (we know the GEVP matrix must be hermitian) by taking the weighted-shifted $C(t) \to \frac{1}{2}\left(C(t)+C(t)^\dagger\right)$. Note: one could also simply compute the matrix to be upper/lower triangular and avoid about half the contractions. As in the case of using g5 hermiticity and aux symmetry, this would reduce our costs substantially, but for our initial runs we want to be sure the basic procedure is correct.

8. We take the eigenvalues and form an effective mass ratio. We then force a fit by minimizing the quadratic $\left(\log(\text{cosh ratio}) - \log\left(\frac{\lambda(t)}{\lambda(t+1)}\right)\right)^2$. This procedure is somehow invariant under $E \to -E$. If this happens, we throw out the negative solution and switch to the positive assuming the residual does not change very much (up to about an order of magnitude in the very small starting residual). If the arg to the log is ever negative, we mark this eigenvalue dimension's time slice with nan and do not fit or plot it. We still are using this dimension in our GEVP, but we do not fit it for this time slice.

9. We fit the result to a constant under the jackknife: We do a different fit for each jackknife sample, computing double elimination jackknife blocks on the fly from our single elimination jackknife blocks to estimate the covariance matrix for each fit. If the user provides any information about time slice-operator combinations to remove (or if they are marked nan), we delete these rows and columns from the resulting covariance matrix (which is itself the tensor product of the time and operator covariances). We then rotate to the correlation matrix basis, invert the result, and rotate back. We then fill in the deleted dimensions with 0's so they will not contribute to χ^2. Recently noticed in a very noisy $I = 0$ fit, it seems that there can be a large difference between using this dynamic pruning procedure and doing the same fit but starting with a reduced fitting window. For this reason, I would say there may be a bug which needs to be fixed (but I need to examine things more closely).

10. These fits are themselves looped over fit ranges (each dimension of the GEVP can have a different fit range). We use a sloppy stopping condition on the χ^2 minimizer for speed. We average the results

over the fit ranges and propagate the underlying jackknife errors to the overall error by computing a covariance matrix on the fit ranges. If our fit window is perfect, in the sense that we are not fitting to noise (not too late in time) and we have only N states in our $N \times N$ GEVP (not too early), then any fit range would give us an estimate of our parameters. If we now pick a fit range which gives us good results, we would be introducing bias in our procedure. We should instead average the results over the fit ranges. In the high statistics limit, these would all give the same result, but in the low statistics limit, the results can vary. If a particular data point gives us a small eigenvalue for our covariance matrix, we would be justified in wanting to eliminate it somehow. Intuitively, the fit range average samples points (roughly) inversely proportionate to how much they contribute to the χ^2's for these fit ranges. As more data is accumulated, we should thus see $> \sqrt{N}$ scaling as more of the fit window becomes useful (this has been observed). For safety, we restrict the fit ranges to include at least 3 time slices from each dimension of the GEVP. The fit range is also forced to be an arithmetic sequence (so that it has late and early times). We skip fit ranges which give bad results (e.g. over/underfit χ^2; there are other errors which cause a skip such as having a relativistic $\gamma < 1$).

11. Once we get the average energy results, we minimize the maximum distance between these averages and the averages of the individual fit ranges. The fit range with minimum maximum we call the representative fit. We redo this representative fit with a stringent stopping condition, and use it to plot the result (probably it would be better to use the numeric results to calculate a systematic error rather than just throwing them away). We display the fit range averages as the numeric values. The error bars come from the diagonal of the average covariance matrix from the representative fit.

Appendix F

Full Numerical Results for Vacuum Subtraction Study

F.1 Don't Subtract (Stationary Bubble)

These results are for

FigureV_sep4_mom1src000_mom2src000_mom1snk000

and should be considered as a control (as we would always do this subtraction in the generalized eigenvalue analysis)

```
t=0 avg: 1.61062122e+10+1.57811720e+06j err: 8.51977589e+07
t=1 avg: 1.60352705e+10+1.50213860e+06j err: 8.46412430e+07
t=2 avg: 1.59922232e+10+6.99733622e+05j err: 8.43058179e+07
t=3 avg: 1.59652665e+10+7.74273107e+05j err: 8.40154511e+07
t=4 avg: 1.59524499e+10+9.25525611e+05j err: 8.37930551e+07
t=5 avg: 1.59494362e+10+5.04999285e+05j err: 8.36779514e+07
t=6 avg: 1.59509739e+10-7.30560966e+05j err: 8.36332334e+07
t=7 avg: 1.59538880e+10-3.87872294e+05j err: 8.36143662e+07
t=8 avg: 1.59569121e+10-5.78324561e+04j err: 8.36247093e+07
t=9 avg: 1.59617972e+10-1.73288064e+05j err: 8.36677762e+07
t=10 avg: 1.59627468e+10-2.69204122e+05j err: 8.36770485e+07
t=11 avg: 1.59667304e+10+2.58049273e+05j err: 8.36503132e+07
t=12 avg: 1.59655605e+10+1.57542214e-09j err: 8.36002312e+07
```

```
t=13 avg: 1.59667304e+10-2.58049273e+05j err: 8.36503132e+07
t=14 avg: 1.59627468e+10+2.69204122e+05j err: 8.36770485e+07
t=15 avg: 1.59617972e+10+1.73288064e+05j err: 8.36677762e+07
t=16 avg: 1.59569121e+10+5.78324561e+04j err: 8.36247093e+07
t=17 avg: 1.59538880e+10+3.87872294e+05j err: 8.36143662e+07
t=18 avg: 1.59509739e+10+7.30560966e+05j err: 8.36332334e+07
t=19 avg: 1.59494362e+10-5.04999285e+05j err: 8.36779514e+07
t=20 avg: 1.59524499e+10-9.25525611e+05j err: 8.37930551e+07
t=21 avg: 1.59652665e+10-7.74273107e+05j err: 8.40154511e+07
t=22 avg: 1.59922232e+10-6.99733622e+05j err: 8.43058179e+07
t=23 avg: 1.60352705e+10-1.50213860e+06j err: 8.46412430e+07
t=24 avg: 1.61062122e+10-1.57811720e+06j err: 8.51977589e+07
t=25 avg: 1.61408043e+10-9.79934399e+05j err: 8.56418743e+07
t=26 avg: 1.62013757e+10-1.22260176e+06j err: 8.62507257e+07
t=27 avg: 1.62709181e+10-7.41717318e+05j err: 8.68353837e+07
t=28 avg: 1.64008283e+10-2.51681968e-10j err: 8.74394856e+07
t=29 avg: 1.62709181e+10+7.41717318e+05j err: 8.68353837e+07
t=30 avg: 1.62013757e+10+1.22260176e+06j err: 8.62507257e+07
t=31 avg: 1.61408043e+10+9.79934399e+05j err: 8.56418743e+07
```

F.2 Taking the Real Part

These results are for

FigureV_sep4_mom1src000_mom2src000_mom1snk000

F.2.1 Time averaged

```
t=0 avg: 1.37778052e+08 err: 6.21737524e+06
t=1 avg: 6.59539950e+07 err: 5.74785509e+06
t=2 avg: 2.31746163e+07 err: 5.56633811e+06
t=3 avg: -4.02211219e+06 err: 5.37480964e+06
t=4 avg: -1.65074321e+07 err: 5.24104839e+06
t=5 avg: -1.98623771e+07 err: 5.26185761e+06
```

```
t=6 avg: -1.82218201e+07 err: 5.27167515e+06
t=7 avg: -1.52902860e+07 err: 5.42751381e+06
t=8 avg: -1.23292400e+07 err: 5.75928747e+06
t=9 avg: -7.24331938e+06 err: 6.34305165e+06
t=10 avg: -6.37140436e+06 err: 7.08319594e+06
t=11 avg: -2.36231584e+06 err: 7.75839340e+06
t=12 avg: -3.52674094e+06 err: 8.11546132e+06
t=13 avg: -2.36231584e+06 err: 7.75839340e+06
t=14 avg: -6.37140436e+06 err: 7.08319594e+06
t=15 avg: -7.24331938e+06 err: 6.34305165e+06
t=16 avg: -1.23292400e+07 err: 5.75928747e+06
t=17 avg: -1.52902860e+07 err: 5.42751381e+06
t=18 avg: -1.82218201e+07 err: 5.27167515e+06
t=19 avg: -1.98623771e+07 err: 5.26185761e+06
t=20 avg: -1.65074321e+07 err: 5.24104839e+06
t=21 avg: -4.02211219e+06 err: 5.37480964e+06
t=22 avg: 2.31746163e+07 err: 5.56633811e+06
t=23 avg: 6.59539950e+07 err: 5.74785509e+06
t=24 avg: 1.37778052e+08 err: 6.21737524e+06
t=25 avg: 1.71625964e+08 err: 6.83420813e+06
t=26 avg: 2.32255736e+08 err: 7.72202157e+06
t=27 avg: 3.01736253e+08 err: 8.49987753e+06
t=28 avg: 4.15419467e+08 err: 9.08578242e+06
t=29 avg: 3.01736253e+08 err: 8.49987753e+06
t=30 avg: 2.32255736e+08 err: 7.72202157e+06
t=31 avg: 1.71625964e+08 err: 6.83420813e+06
```

F.2.2 No time averaging

```
t=0 avg: 1.36682012e+08 err: 6.17367769e+06
t=1 avg: 6.54542910e+07 err: 5.72484373e+06
t=2 avg: 2.31226959e+07 err: 5.54775467e+06
t=3 avg: -3.87654450e+06 err: 5.37401424e+06
```

```
t=4 avg: -1.61640940e+07 err: 5.25113689e+06
t=5 avg: -1.91144711e+07 err: 5.26672401e+06
t=6 avg: -1.71322358e+07 err: 5.25409589e+06
t=7 avg: -1.41120229e+07 err: 5.38330376e+06
t=8 avg: -1.11039998e+07 err: 5.68434063e+06
t=9 avg: -5.94886372e+06 err: 6.24069234e+06
t=10 avg: -5.11769158e+06 err: 6.95993564e+06
t=11 avg: -1.35798965e+06 err: 7.62506280e+06
t=12 avg: -2.61087441e+06 err: 7.97728154e+06
t=13 avg: -1.35798965e+06 err: 7.62506280e+06
t=14 avg: -5.11769158e+06 err: 6.95993564e+06
t=15 avg: -5.94886372e+06 err: 6.24069234e+06
t=16 avg: -1.11039998e+07 err: 5.68434063e+06
t=17 avg: -1.41120229e+07 err: 5.38330376e+06
t=18 avg: -1.71322358e+07 err: 5.25409589e+06
t=19 avg: -1.91144711e+07 err: 5.26672401e+06
t=20 avg: -1.61640940e+07 err: 5.25113689e+06
t=21 avg: -3.87654450e+06 err: 5.37401424e+06
t=22 avg: 2.31226959e+07 err: 5.54775467e+06
t=23 avg: 6.54542910e+07 err: 5.72484373e+06
t=24 avg: 1.36682012e+08 err: 6.17367769e+06
t=25 avg: 1.70156833e+08 err: 6.78827588e+06
t=26 avg: 2.30377182e+08 err: 7.68024678e+06
t=27 avg: 2.99360470e+08 err: 8.43986069e+06
t=28 avg: 4.12681078e+08 err: 9.01352919e+06
t=29 avg: 2.99360470e+08 err: 8.43986069e+06
t=30 avg: 2.30377182e+08 err: 7.68024678e+06
t=31 avg: 1.70156833e+08 err: 6.78827588e+06
```

F.3 Don't Take the Real

These results are for

FigureV_sep4_mom1src000_mom2src000_mom1snk000

F.3.1 Time averaged

```
t=0 avg: 1.37113601e+08+1.57811720e+06j err: 6.22149479e+06
t=1 avg: 6.61718637e+07+1.50213860e+06j err: 5.75406839e+06
t=2 avg: 2.31246342e+07+6.99733622e+05j err: 5.56193329e+06
t=3 avg: -3.83207428e+06+7.74273107e+05j err: 5.37214765e+06
t=4 avg: -1.66487104e+07+9.25525611e+05j err: 5.23706829e+06
t=5 avg: -1.96624059e+07+5.04999285e+05j err: 5.26001814e+06
t=6 avg: -1.81246827e+07-7.30560966e+05j err: 5.27336237e+06
t=7 avg: -1.52105824e+07-3.87872294e+05j err: 5.43023667e+06
t=8 avg: -1.21864533e+07-5.78324561e+04j err: 5.75724619e+06
t=9 avg: -7.30141131e+06-1.73288064e+05j err: 6.34648814e+06
t=10 avg: -6.35177791e+06-2.69204122e+05j err: 7.08460273e+06
t=11 avg: -2.36816833e+06+2.58049273e+05j err: 7.76870201e+06
t=12 avg: -3.53813486e+06+3.10871403e-12j err: 8.12075753e+06
t=13 avg: -2.36816833e+06-2.58049273e+05j err: 7.76870201e+06
t=14 avg: -6.35177791e+06+2.69204122e+05j err: 7.08460273e+06
t=15 avg: -7.30141131e+06+1.73288064e+05j err: 6.34648814e+06
t=16 avg: -1.21864533e+07+5.78324561e+04j err: 5.75724619e+06
t=17 avg: -1.52105824e+07+3.87872294e+05j err: 5.43023667e+06
t=18 avg: -1.81246827e+07+7.30560966e+05j err: 5.27336237e+06
t=19 avg: -1.96624059e+07-5.04999285e+05j err: 5.26001814e+06
t=20 avg: -1.66487104e+07-9.25525611e+05j err: 5.23706829e+06
t=21 avg: -3.83207428e+06-7.74273107e+05j err: 5.37214765e+06
t=22 avg: 2.31246342e+07-6.99733622e+05j err: 5.56193329e+06
t=23 avg: 6.61718637e+07-1.50213860e+06j err: 5.75406839e+06
t=24 avg: 1.37113601e+08-1.57811720e+06j err: 6.22149479e+06
t=25 avg: 1.71705661e+08-9.79934399e+05j err: 6.82702644e+06
t=26 avg: 2.32277130e+08-1.22260176e+06j err: 7.72794171e+06
t=27 avg: 3.01819482e+08-7.41717318e+05j err: 8.50394436e+06
t=28 avg: 4.31729735e+08+2.17641634e-12j err: 9.08434782e+06
t=29 avg: 3.01819482e+08+7.41717318e+05j err: 8.50394436e+06
t=30 avg: 2.32277130e+08+1.22260176e+06j err: 7.72794171e+06
```

```
t=31 avg: 1.71705661e+08+9.79934399e+05j err: 6.82702644e+06
```

F.3.2 No time averaging

```
t=0 avg: 1.36008035e+08+1.62438247e+06j err: 6.17782061e+06
t=1 avg: 6.56788327e+07+1.52230521e+06j err: 5.73011783e+06
t=2 avg: 2.30788708e+07+7.78785004e+05j err: 5.54359362e+06
t=3 avg: -3.69430167e+06+9.21654272e+05j err: 5.37041539e+06
t=4 avg: -1.63072694e+07+1.01437532e+06j err: 5.24734901e+06
t=5 avg: -1.89048037e+07+5.92500921e+05j err: 5.26531714e+06
t=6 avg: -1.70279220e+07-6.27362618e+05j err: 5.25662677e+06
t=7 avg: -1.40337605e+07-2.87978393e+05j err: 5.38635752e+06
t=8 avg: -1.09707243e+07-3.36993962e+04j err: 5.68293973e+06
t=9 avg: -6.00524783e+06-1.68820886e+05j err: 6.24474647e+06
t=10 avg: -5.10326552e+06-2.51971915e+05j err: 6.96140892e+06
t=11 avg: -1.35496252e+06+2.97698602e+05j err: 7.63469967e+06
t=12 avg: -2.62327899e+06+2.18775511e-11j err: 7.98380689e+06
t=13 avg: -1.35496252e+06-2.97698602e+05j err: 7.63469967e+06
t=14 avg: -5.10326552e+06+2.51971915e+05j err: 6.96140892e+06
t=15 avg: -6.00524783e+06+1.68820886e+05j err: 6.24474647e+06
t=16 avg: -1.09707243e+07+3.36993962e+04j err: 5.68293973e+06
t=17 avg: -1.40337605e+07+2.87978393e+05j err: 5.38635752e+06
t=18 avg: -1.70279220e+07+6.27362618e+05j err: 5.25662677e+06
t=19 avg: -1.89048037e+07-5.92500921e+05j err: 5.26531714e+06
t=20 avg: -1.63072694e+07-1.01437532e+06j err: 5.24734901e+06
t=21 avg: -3.69430167e+06-9.21654272e+05j err: 5.37041539e+06
t=22 avg: 2.30788708e+07-7.78785004e+05j err: 5.54359362e+06
t=23 avg: 6.56788327e+07-1.52230521e+06j err: 5.73011783e+06
t=24 avg: 1.36008035e+08-1.62438247e+06j err: 6.17782061e+06
t=25 avg: 1.70242977e+08-1.10452556e+06j err: 6.78112241e+06
t=26 avg: 2.30402939e+08-1.26715334e+06j err: 7.68666687e+06
t=27 avg: 2.99448847e+08-7.13889599e+05j err: 8.44358404e+06
t=28 avg: 4.28950601e+08+3.53262862e-12j err: 9.01372406e+06
```

t=29 avg: 2.99448847e+08+7.13889599e+05j err: 8.44358404e+06

t=30 avg: 2.30402939e+08+1.26715334e+06j err: 7.68666687e+06

t=31 avg: 1.70242977e+08+1.10452556e+06j err: 6.78112241e+06

F.4 Don't Subtract (Moving Bubbles)

These results are for the diagram:

FigureV_sep4_mom1src001_mom2src000_mom1snk001

t=0 avg: 6.49955294e+07−2.67303818e+04j err: 1.29027661e+06

t=1 avg: 4.00007690e+07+3.60111914e+05j err: 1.13335346e+06

t=2 avg: 2.37802121e+07+7.00691224e+05j err: 1.02982784e+06

t=3 avg: 1.31907639e+07+1.12396163e+06j err: 9.97156339e+05

t=4 avg: 6.59375689e+06+9.50554846e+05j err: 9.98358544e+05

t=5 avg: 3.02893403e+06+1.14230040e+06j err: 1.02478350e+06

t=6 avg: 1.16438829e+06+7.66290954e+05j err: 1.04108780e+06

t=7 avg: 9.44378615e+04+6.28411320e+05j err: 1.02268883e+06

t=8 avg: −5.82190191e+05+1.53068844e+05j err: 9.73668414e+05

t=9 avg: −8.32145956e+05−6.34942361e+05j err: 1.00715896e+06

t=10 avg: −9.92327595e+05−8.96386157e+05j err: 1.03794205e+06

t=11 avg: −1.26376138e+06−6.21084770e+05j err: 1.07888049e+06

t=12 avg: −2.25244778e+06−1.43163198e+05j err: 1.13620999e+06

t=13 avg: −2.56146008e+06+1.10438274e+06j err: 1.18528663e+06

t=14 avg: −2.76141053e+06+1.23475437e+06j err: 1.19403349e+06

t=15 avg: −3.17444613e+06+5.22126217e+05j err: 1.14675309e+06

t=16 avg: −3.63706308e+06−4.20668009e+04j err: 1.15202541e+06

t=17 avg: −2.78171648e+06+3.63656734e+05j err: 1.10702862e+06

t=18 avg: −1.96152732e+06−3.89170391e+05j err: 1.06409103e+06

t=19 avg: −1.55082156e+06−1.14565510e+06j err: 1.04746545e+06

t=20 avg: −1.14646154e+05−1.89546405e+06j err: 1.07480461e+06

t=21 avg: 6.18448753e+05−1.73824861e+06j err: 1.04487118e+06

t=22 avg: 2.17245332e+06−2.36503212e+06j err: 1.03547245e+06

t=23 avg: 4.35115432e+06−2.75374606e+06j err: 1.02290847e+06

t=24 avg: 9.42732427e+06−2.67533779e+06j err: 1.04363231e+06

t=25 avg: 1.50590284e+07−1.69645856e+06j err: 1.05363565e+06

t=26 avg: 2.54151591e+07−5.34975267e+05j err: 1.12534390e+06

t=27 avg: 3.95587298e+07+1.08256184e+05j err: 1.26373203e+06

t=28 avg: 6.12525495e+07+1.00723292e+06j err: 1.39687143e+06

t=29 avg: 5.72321001e+07+1.34214939e+06j err: 1.50464115e+06

t=30 avg: 5.87856391e+07+4.79864072e+05j err: 1.48276733e+06

t=31 avg: 5.95168610e+07+5.95272906e+05j err: 1.42601482e+06

F.5 Do Subtract Moving Bubbles

These results are for the diagram:

FigureV_sep4_mom1src001_mom2src000_mom1snk001

F.5.1 Time Averaged

t=0 avg: 6.45713459e+07−5.68619062e+04j err: 1.28082499e+06

t=1 avg: 3.95765856e+07+3.29980390e+05j err: 1.12111858e+06

t=2 avg: 2.33560286e+07+6.70559700e+05j err: 1.01921366e+06

t=3 avg: 1.27665804e+07+1.09383011e+06j err: 9.87392507e+05

t=4 avg: 6.16957341e+06+9.20423321e+05j err: 9.93106993e+05

t=5 avg: 2.60475056e+06+1.11216887e+06j err: 1.02215885e+06

t=6 avg: 7.40204814e+05+7.36159429e+05j err: 1.04117854e+06

t=7 avg: −3.29745615e+05+5.98279796e+05j err: 1.02608170e+06

t=8 avg: −1.00637367e+06+1.22937320e+05j err: 9.78230713e+05

t=9 avg: −1.25632943e+06−6.65073885e+05j err: 1.01191958e+06

t=10 avg: −1.41651107e+06−9.26517681e+05j err: 1.03820054e+06

t=11 avg: −1.68794486e+06−6.51216295e+05j err: 1.08709834e+06

t=12 avg: −2.67663125e+06−1.73294723e+05j err: 1.14053053e+06

t=13 avg: −2.98564355e+06+1.07425121e+06j err: 1.18913586e+06

t=14 avg: −3.18559401e+06+1.20462284e+06j err: 1.19440218e+06

t=15 avg: −3.59862961e+06+4.91994693e+05j err: 1.15041351e+06

t=16 avg: −4.06124655e+06−7.21983253e+04j err: 1.15934115e+06

t=17 avg: −3.20589995e+06+3.33525210e+05j err: 1.11328619e+06

```
t=18 avg: -2.38571080e+06-4.19301916e+05j err: 1.07021209e+06
t=19 avg: -1.97500504e+06-1.17578663e+06j err: 1.05081972e+06
t=20 avg: -5.38829631e+05-1.92559557e+06j err: 1.07116452e+06
t=21 avg: 1.94265277e+05-1.76838014e+06j err: 1.03723682e+06
t=22 avg: 1.74826984e+06-2.39516364e+06j err: 1.02757642e+06
t=23 avg: 3.92697085e+06-2.78387759e+06j err: 1.01349427e+06
t=24 avg: 9.00314080e+06-2.70546932e+06j err: 1.03711828e+06
t=25 avg: 1.46348449e+07-1.72659008e+06j err: 1.04811038e+06
t=26 avg: 2.49909756e+07-5.65106791e+05j err: 1.11744536e+06
t=27 avg: 3.91345463e+07+7.81246593e+04j err: 1.25264376e+06
t=28 avg: 6.08283660e+07+9.77101396e+05j err: 1.38978528e+06
t=29 avg: 5.68079166e+07+1.31201786e+06j err: 1.49439012e+06
t=30 avg: 5.83614556e+07+4.49732547e+05j err: 1.47314927e+06
t=31 avg: 5.90926775e+07+5.65141382e+05j err: 1.41715364e+06
```

F.5.2 No Time Averaging

```
t=0 avg: 6.37797053e+07-2.04693608e+04j err: 1.26331173e+06
t=1 avg: 3.89403420e+07+4.35682582e+05j err: 1.10580482e+06
t=2 avg: 2.29565481e+07+8.08017497e+05j err: 1.00905800e+06
t=3 avg: 1.25658288e+07+1.27176429e+06j err: 9.81040781e+05
t=4 avg: 6.12733209e+06+1.12555220e+06j err: 9.92362827e+05
t=5 avg: 2.70036391e+06+1.34220426e+06j err: 1.02246414e+06
t=6 avg: 8.99195087e+05+1.00252703e+06j err: 1.03918151e+06
t=7 avg: -1.37061612e+05+8.96596205e+05j err: 1.01418932e+06
t=8 avg: -7.65914215e+05+4.78654747e+05j err: 9.65719599e+05
t=9 avg: -1.04616214e+06-2.42022662e+05j err: 1.00149429e+06
t=10 avg: -1.27094174e+06-4.94063796e+05j err: 1.03218854e+06
t=11 avg: -1.54442885e+06-2.32165565e+05j err: 1.08303482e+06
t=12 avg: -2.52614073e+06+1.52417722e+05j err: 1.12992833e+06
t=13 avg: -2.86316596e+06+1.32933596e+06j err: 1.18219874e+06
t=14 avg: -3.06583741e+06+1.32675797e+06j err: 1.18520061e+06
t=15 avg: -3.45152039e+06+4.69222228e+05j err: 1.13822063e+06
```

```
t=16 avg: -3.89025914e+06-2.23146239e+05j err: 1.14118579e+06
t=17 avg: -2.99948585e+06+4.13963919e+04j err: 1.09839495e+06
t=18 avg: -2.08819871e+06-7.25009033e+05j err: 1.05479819e+06
t=19 avg: -1.61627888e+06-1.49074645e+06j err: 1.03867149e+06
t=20 avg: -1.52020625e+05-2.22372118e+06j err: 1.06004670e+06
t=21 avg: 5.87722119e+05-2.03463847e+06j err: 1.02827894e+06
t=22 avg: 2.11317365e+06-2.61669049e+06j err: 1.02306009e+06
t=23 avg: 4.23797358e+06-3.04345900e+06j err: 1.00799642e+06
t=24 avg: 9.27787066e+06-2.93190041e+06j err: 1.03184355e+06
t=25 avg: 1.48150304e+07-1.93452058e+06j err: 1.04284057e+06
t=26 avg: 2.50127778e+07-7.96380499e+05j err: 1.10941531e+06
t=27 avg: 3.89897425e+07-1.62973089e+05j err: 1.24356389e+06
t=28 avg: 6.04551123e+07+7.17459155e+05j err: 1.37537841e+06
t=29 avg: 5.62348627e+07+1.06118404e+06j err: 1.47317853e+06
t=30 avg: 5.76267712e+07+2.74935337e+05j err: 1.45142579e+06
t=31 avg: 5.82954678e+07+4.98615465e+05j err: 1.39341786e+06
```

Appendix G

zMöbius Coefficients for 24^3

We list zMobius coefficients (b's and c's) and corresponding omega's for an $Ls = 12$ approximation of Shamir $Ls = 24$ for the lattice ensemble from section 3.2.

```
Array zmobius_b_coeff[24] = {
double zmobius_b_coeff[0] = 0.9585786062245000
double zmobius_b_coeff[1] = 0.0000000000000000
double zmobius_b_coeff[2] = 1.0224505516837099
double zmobius_b_coeff[3] = 0.0000000000000000
double zmobius_b_coeff[4] = 1.2093319385452017
double zmobius_b_coeff[5] = 0.0000000000000000
double zmobius_b_coeff[6] = 1.5208264566393539
double zmobius_b_coeff[7] = 0.0000000000000000
double zmobius_b_coeff[8] = 2.0215683321339464
double zmobius_b_coeff[9] = 0.0000000000000000
double zmobius_b_coeff[10] = 2.8079502253536113
double zmobius_b_coeff[11] = 0.0000000000000000
double zmobius_b_coeff[12] = 4.0407974583803377
double zmobius_b_coeff[13] = 0.0000000000000000
double zmobius_b_coeff[14] = 6.0079114984238400
double zmobius_b_coeff[15] = 0.0000000000000000
double zmobius_b_coeff[16] = 9.2571848022742405
double zmobius_b_coeff[17] = 1.1769029217606854
```

```
double zmobius_b_coeff[18] = 9.2571848022742405
double zmobius_b_coeff[19] = -1.1769029217606854
double zmobius_b_coeff[20] = 7.9472247379278818
double zmobius_b_coeff[21] = 6.8186297292500981
double zmobius_b_coeff[22] = 7.9472247379278818
double zmobius_b_coeff[23] = -6.8186297292500981
}
Array zmobius_c_coeff[24] = {
double zmobius_c_coeff[0] = -0.0414213937755000
double zmobius_c_coeff[1] = 0.0000000000000000
double zmobius_c_coeff[2] = 0.0224505516837099
double zmobius_c_coeff[3] = 0.0000000000000000
double zmobius_c_coeff[4] = 0.2093319385452017
double zmobius_c_coeff[5] = 0.0000000000000000
double zmobius_c_coeff[6] = 0.5208264566393539
double zmobius_c_coeff[7] = 0.0000000000000000
double zmobius_c_coeff[8] = 1.0215683321339464
double zmobius_c_coeff[9] = 0.0000000000000000
double zmobius_c_coeff[10] = 1.8079502253536113
double zmobius_c_coeff[11] = 0.0000000000000000
double zmobius_c_coeff[12] = 3.0407974583803377
double zmobius_c_coeff[13] = 0.0000000000000000
double zmobius_c_coeff[14] = 5.0079114984238400
double zmobius_c_coeff[15] = 0.0000000000000000
double zmobius_c_coeff[16] = 8.2571848022742405
double zmobius_c_coeff[17] = 1.1769029217606854
double zmobius_c_coeff[18] = 8.2571848022742405
double zmobius_c_coeff[19] = -1.1769029217606854
double zmobius_c_coeff[20] = 6.9472247379278818
double zmobius_c_coeff[21] = 6.8186297292500981
double zmobius_c_coeff[22] = 6.9472247379278818
double zmobius_c_coeff[23] = -6.8186297292500981
```

}

omega[0]=1.0903256131299373 0

omega[1]=0.9570283702230611 0

omega[2]=0.7048886040934104 0

omega[3]=0.48979921782791747 0

omega[4]=0.328608311201356 0

omega[5]=0.21664245377015995 0

omega[6]=0.14121112711957107 0

omega[7]=0.0907785101745156 0

omega[8]=0.05608303440064219 −0.007537158177840385

omega[9]=0.05608303440064219 0.007537158177840385

omega[10]=0.0365221637144842 −0.03343945161367745

omega[11]=0.0365221637144842 0.03343945161367745

Bibliography

[1] Jozef J. Dudek, Robert G. Edwards, and Christopher E. Thomas. "S and D-wave phase shifts in isospin-2 pi pi scattering from lattice QCD". In: *Phys. Rev.* D86 (2012), p. 034031. DOI: 10.1103/PhysRevD.86.034031. arXiv: 1203.6041 [hep-ph].

[2] T. Yamazaki et al. "I = 2 pi pi scattering phase shift with two flavors of O(a) improved dynamical quarks". In: *Phys. Rev.* D70 (2004), p. 074513. DOI: 10.1103/PhysRevD.70.074513. arXiv: hep-lat/0402025 [hep-lat].

[3] Jozef J. Dudek, Robert G. Edwards, and Christopher E. Thomas. "Energy dependence of the ρ resonance in $\pi\pi$ elastic scattering from lattice QCD". In: *Phys. Rev.* D87.3 (2013). [Erratum: Phys. Rev.D90,no.9,099902(2014)], p. 034505. DOI: 10.1103/PhysRevD.87.034505, 10.1103/PhysRevD.90.099902. arXiv: 1212.0830 [hep-ph].

[4] Constantia Alexandrou et al. "P-wave $\pi\pi$ scattering and the ρ resonance from lattice QCD". In: *Phys. Rev.* D96.3 (2017), p. 034525. DOI: 10.1103/PhysRevD.96.034525. arXiv: 1704.05439 [hep-lat].

[5] Raul A. Briceno et al. "Isoscalar $\pi\pi$ scattering and the σ meson resonance from QCD". In: *Phys. Rev. Lett.* 118.2 (2017), p. 022002. DOI: 10.1103/PhysRevLett.118.022002. arXiv: 1607.05900 [hep-ph].

[6] HEINRICH WAHL. "FIRST OBSERVATION AND MEASUREMENT OF DIRECT CP VIOLATION BY THE NA31 AND NA48 EXPERIMENTS AT CERN". In: *Basics and Highlights in Fundamental Physics*. WORLD SCIENTIFIC, Apr. 2001. DOI: 10.1142/9789812811585_0013. URL: https://doi.org/10.1142/9789812811585_0013.

[7] Q. Liu. "Kaon to Two Pions decays from Lattice QCD: Delta I = 1/2 rule and CP violation". PhD thesis. Columbia University, 2012.

[8] Justin Foley et al. "Practical all-to-all propagators for lattice QCD". In: *Comput. Phys. Commun.* 172 (2005), pp. 145–162. DOI: 10.1016/j.cpc.2005.06.008. arXiv: hep-lat/0505023 [hep-lat].

[9] R. J. Hudspith. “Fourier Accelerated Conjugate Gradient Lattice Gauge Fixing”. In: *Comput. Phys. Commun.* 187 (2015), pp. 115–119. DOI: 10.1016/j.cpc.2014.10.017. arXiv: 1405.5812 [hep-lat].

[10] Christopher E. Thomas, Robert G. Edwards, and Jozef J. Dudek. “Helicity operators for mesons in flight on the lattice”. In: *Phys. Rev.* D85 (2012), p. 014507. DOI: 10.1103/PhysRevD.85.014507, 10.1103/PhysRevD.85.039901. arXiv: 1107.1930 [hep-lat].

[11] Martin Luscher. “Two particle states on a torus and their relation to the scattering matrix”. In: *Nucl. Phys.* B354 (1991), pp. 531–578. DOI: 10.1016/0550-3213(91)90366-6.

[12] Luka Leskovec and Sasa Prelovsek. “Scattering phase shifts for two particles of different mass and non-zero total momentum in lattice QCD”. In: *Phys. Rev.* D85 (2012), p. 114507. DOI: 10.1103/PhysRevD.85.114507. arXiv: 1202.2145 [hep-lat].

[13] P. A. Boyle et al. “Low energy constants of SU(2) partially quenched chiral perturbation theory from N_f=2+1 domain wall QCD”. In: *Phys. Rev.* D93.5 (2016), p. 054502. DOI: 10.1103/PhysRevD.93.054502. arXiv: 1511.01950 [hep-lat].

[14] G. Colangelo, J. Gasser, and H. Leutwyler. “$\pi\pi$ scattering”. In: *Nucl. Phys.* B603 (2001), pp. 125–179. DOI: 10.1016/S0550-3213(01)00147-X. arXiv: hep-ph/0103088 [hep-ph].

[15] Daniel Hoying. *Lattice Fitter*. https://github.com/goracle/lattice-fitter. 2015-2017.

[16] M. Luscher. “Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories. 1. Stable Particle States”. In: *Commun. Math. Phys.* 104 (1986), p. 177. DOI: 10.1007/BF01211589.

[17] M. Luscher. “Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories. 2. Scattering States”. In: *Commun. Math. Phys.* 105 (1986), pp. 153–188. DOI: 10.1007/BF01211097.

[18] C. h. Kim, C. T. Sachrajda, and Stephen R. Sharpe. “Finite-volume effects for two-hadron states in moving frames”. In: *Nucl. Phys.* B727 (2005), pp. 218–243. DOI: 10.1016/j.nuclphysb.2005.08.029. arXiv: hep-lat/0507006 [hep-lat].

[19] Norman H. Christ, Changhoan Kim, and Takeshi Yamazaki. “Finite volume corrections to the two-particle decay of states with non-zero momentum”. In: *Phys. Rev.* D72 (2005), p. 114506. DOI: 10.1103/PhysRevD.72.114506. arXiv: hep-lat/0507009 [hep-lat].

[20] T. Yamazaki et al. “I = 2 pi pi scattering phase shift with two flavors of O(a) improved dynamical quarks”. In: *Phys. Rev.* D70 (2004), p. 074513. DOI: 10.1103/PhysRevD.70.074513. arXiv: hep-lat/0402025 [hep-lat].

[21] David B. Kaplan. “A Method for simulating chiral fermions on the lattice”. In: *Phys. Lett.* B288 (1992), pp. 342–347. DOI: 10.1016/0370-2693(92)91112-M. arXiv: hep-lat/9206013 [hep-lat].

[22] Yigal Shamir. “Chiral fermions from lattice boundaries”. In: *Nucl. Phys.* B406 (1993), pp. 90–106. DOI: 10.1016/0550-3213(93)90162-I. arXiv: hep-lat/9303005 [hep-lat].

[23] Holger Bech Nielsen and M. Ninomiya. “No Go Theorem for Regularizing Chiral Fermions”. In: *Phys. Lett.* 105B (1981), pp. 219–223. DOI: 10.1016/0370-2693(81)91026-1.

[24] D. Friedan. “A PROOF OF THE NIELSEN-NINOMIYA THEOREM”. In: *Commun. Math. Phys.* 85 (1982), pp. 481–490. DOI: 10.1007/BF01403500.

[25] Richard C. Brower, Harmut Neff, and Kostas Orginos. “The Möbius domain wall fermion algorithm”. In: *Comput. Phys. Commun.* 220 (2017), pp. 1–19. DOI: 10.1016/j.cpc.2017.01.024. arXiv: 1206.5214 [hep-lat].

[26] Taku Izubuchi et al. “Calculation of Nucleon Electric Dipole Moments Induced by Quark Chromo-Electric Dipole Moments”. In: *PoS* LATTICE2016 (2017), p. 398. DOI: 10.22323/1.256.0398. arXiv: 1702.00052 [hep-lat].

[27] Hantao Yin and Robert D. Mawhinney. “Improving DWF Simulations: the Force Gradient Integrator and the Möbius Accelerated DWF Solver”. In: *PoS* LATTICE2011 (2011), p. 051. DOI: 10.22323/1.139.0051. arXiv: 1111.5059 [hep-lat].

[28] Richard M. Stallman, Roland McGrath, and Paul D. Smith. *GNU Make: A Program for Directing Recompilation, for Version 3.81*. Free Software Foundation, 2004. ISBN: 1882114833.

[29] Eigo Shintani et al. “Covariant approximation averaging”. In: *Phys. Rev.* D91.11 (2015), p. 114511. DOI: 10.1103/PhysRevD.91.114511. arXiv: 1402.0244 [hep-lat].

[30] Daiqian Zhang. “Kaon to Two Pion decay from Lattice QCD and CP violation”. PhD thesis. 2015. DOI: 10.7916/d8c828ns. URL: https://academiccommons.columbia.edu/doi/10.7916/D8C828NS.

[31] Peter Boyle et al. “Grid: A next generation data parallel C++ QCD library”. In: (2015). arXiv: 1512.03487 [hep-lat].

[32] Daiqian Zhang. *Implementation of All to All propagator on BGQ machine*. Private Communication. 2014.

[33] Dwight Renfrew et al. “Controlling Residual Chiral Symmetry Breaking in Domain Wall Fermion Simulations”. In: *PoS* LATTICE2008 (2008), p. 048. DOI: 10.22323/1.066.0048. arXiv: 0902.2587 [hep-lat].

[34] Christopher Kelly. “Continuum Results for Light Hadronic Quantities using Domain Wall Fermions with the Iwasaki and DSDR Gauge Actions”. In: *PoS* LATTICE2011 (2011), p. 285. DOI: 10.22323/1.139.0285. arXiv: 1201.0706 [hep-lat].

[35] Benoit Blossier et al. “On the generalized eigenvalue method for energies and matrix elements in lattice field theory”. In: *JHEP* 04 (2009), p. 094. DOI: 10.1088/1126-6708/2009/04/094. arXiv: 0902.1265 [hep-lat].

[36] Martin Luscher and Ulli Wolff. “How to Calculate the Elastic Scattering Matrix in Two-dimensional Quantum Field Theories by Numerical Simulation”. In: *Nucl. Phys.* B339 (1990), pp. 222–252. DOI: 10.1016/0550-3213(90)90540-T.

[37] John Bulava, Michael Donnellan, and Rainer Sommer. “On the computation of hadron-to-hadron transition matrix elements in lattice QCD”. In: *JHEP* 01 (2012), p. 140. DOI: 10.1007/JHEP01(2012)140. arXiv: 1108.3774 [hep-lat].

[38] Jozef J. Dudek, Robert G. Edwards, and Christopher E. Thomas. “S and D-wave phase shifts in isospin-2 pi pi scattering from lattice QCD”. In: *Phys. Rev.* D86 (2012), p. 034031. DOI: 10.1103/PhysRevD.86.034031. arXiv: 1203.6041 [hep-ph].

[39] Xu Feng. “Investigating scattering phenomena from Lattice QCD using twisted mass fermions”. In: (Jan. 2010).

[40] David Murphy. *Notes on $SU(2)$ Global Fits Including $I = 2$ $\pi\pi$ Scattering Data*. Private Communication. 2018.

[41] Daniel Zwillinger and Stephen Kokoska. *CRC Standard Probability and Statistics Tables and Formulae*. CRC Press, 1999. ISBN: 1584880597. URL: https://www.amazon.com/Standard-Probability-Statistics-Tables-Formulae/dp/1584880597?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=1584880597.

[42] M. Aickin and H. Gensler. “Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods”. In: *Am J Public Health* 86.5 (May 1996), pp. 726–728.

[43] Travis Oliphant. *Guide to NumPy*. Jan. 2006.

[44] John D. Hunter. “Matplotlib: A 2D Graphics Environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/mcse.2007.55. URL: https://doi.org/10.1109/mcse.2007.55.

[45] Andrew Collette. *Python and HDF5*. O’Reilly, 2013.

[46] Aaron Meurer et al. “SymPy: symbolic computing in Python”. In: *PeerJ Computer Science* 3 (Jan. 2017), e103. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.103. URL: https://doi.org/10.7717/peerj-cs.103.

[47] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. [Online; accessed <today>]. 2001–. URL: http://www.scipy.org/.

[48] G. Peter Lepage. *gvar*. https://github.com/gplepage/gvar. 2008-2018.

[49] Per A. Brodtkorb. *Numdifftools*. https://pypi.org/project/Numdifftools/. 2019.

[50] iminuit team. *iminuit – A Python interface to Minuit*. https://github.com/iminuit/iminuit. Accessed: 2018-03-05.

[51] F. James and M. Roos. "Minuit – a system for function minimization and analysis of the parameter errors and correlations". In: *Computer Physics Communications* 10 (Dec. 1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[52] K. Rummukainen and Steven A. Gottlieb. "Resonance scattering phase shifts on a nonrest frame lattice". In: *Nucl. Phys.* B450 (1995), pp. 397–436. DOI: 10.1016/0550-3213(95)00313-H. arXiv: hep-lat/9503028 [hep-lat].

[53] Peter Boyle et al. *goracle/Grid: Pipi Periodic Working Copy*. 2019. DOI: 10.5281/zenodo.2596641. URL: https://zenodo.org/record/2596641.

[54] Dan Hoying. *Lattice Fitter*. https://github.com/goracle/lattice-fitter. 2018.